Krylov and Safonov Estimates for Degenerate Quasilinear Elliptic PDEs

François Delarue [a],

[a] *Laboratoire de Probabilités et Modèles Aléatoires, Université Paris 7-Diderot, Case 7012, 2, Place Jussieu, 75251 Paris Cedex 05, France*

Abstract

We here establish an a priori Hölder estimate of Krylov and Safonov type for the viscosity solutions of a degenerate quasilinear elliptic PDE of non-divergence form. The diffusion matrix may degenerate when the norm of the gradient of the solution is small: the exhibited Hölder exponent and Hölder constant only depend on the growth of the source term and on the bounds of the spectrum of the diffusion matrix for large values of the gradient. In particular, the given estimate is independent of the regularity of the coefficients. As in the original paper by Krylov and Safonov, the proof relies on a probabilistic interpretation of the equation.

Key words: Quasilinear elliptic PDE; Degeneracy; p-Laplacian; Hölder estimate; Stochastic differential equation

1 Introduction

Background and Objective. The original Krylov and Safonov result (see [17, 18]) says that, given two open balls $B_1 \subset B_2 \subset \mathbb{R}^d$ of same center and of radii 1 and 2 and given a solution u in $\mathcal{C}(\overline{B}_2) \cap W^{2,d}_{\mathrm{loc}}(B_2)$ of an elliptic equation of non-divergence type

$$-\mathrm{Tr}(A(x)D^2u(x)) + \langle b(x), Du(x)\rangle + f(x) = 0 \quad \text{for a.e. } x \in B_2,$$

A, b and f being bounded and measurable and A being also uniformly elliptic, u fulfills on the ball $\overline{B}_1$ a universal Hölder estimate whose exponent and constant only depend on the dimension d, on the upper bounds of A, b and

Email address: `delarue@math.jussieu.fr` (François Delarue).
URL: `www.math.jussieu.fr/~delarue` (François Delarue).

f, on the lower bound of A and on the supremum norm of u on $\overline{B}_2$. Obtained in the late 70's, this result may be seen as the counterpart for equations of non-divergence form of the older De Giorgi and Nash estimates (see [9, 22]) established in the 50's for the solutions of divergence form equations.

In a series of works, due among others to Serrin [25] and Ladyzhenskaya et al. [19], De Giorgi and Nash estimates have been shown to hold for quasilinear elliptic equations of divergence type admitting degeneracies of p-Laplace type, $p > 2$, that is for equations driven by the p-Laplace operator $\Delta_p(u) = \mathrm{div}(|Du|^{p-2}Du)$, or, more generally, by a second order operator of the form $\mathrm{div}(A(x, Du))$, the growth of $A(x, Du)$ being controlled from above and from below by $|Du|^{p-1}$, again with $p > 2$. The purpose of this article is to prove a similar result for quasilinear elliptic equations of non-divergence form. Precisely, the main result is

Theorem 1.1 *Let $A : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathcal{S}_d(\mathbb{R})$ (set of symmetric matrices of size d) and $f : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ be continuous coefficients satisfying*[1] *:*

$$\forall y \in \mathbb{R},\ \forall x, z, \xi \in \mathbb{R}^d,\ \begin{cases} \Lambda^{-1}\lambda(z)|\xi|^2 \le \langle \xi, A(x,y,z)\xi\rangle \le \Lambda\lambda(z)|\xi|^2, \\ |f(x,y,z)| \le (1/2)\Lambda(1+\lambda(z))(1+|z|), \end{cases} \tag{H1}$$

for some $\Lambda \ge 1$ and some continuous mapping $\lambda : \mathbb{R}^d \to \mathbb{R}_+$ for which there exist positive reals λ_0 and M (positive means in $(0, +\infty)$) such that $\lambda(z) \ge \lambda_0$ for $|z| \ge M$. Let B_2 be a ball of $\mathbb{R}^d$ of radius 2 and $u : \overline{B}_2 \to \mathbb{R}$ be a bounded and continuous viscosity solution of

$$-\mathrm{Tr}(A(x, u(x), Du(x))D^2u(x)) + f(x, u(x), Du(x)) = 0, \quad x \in B_2. \tag{1.1}$$

Then, u is Hölder continuous on $\overline{B}_1$. Moreover, there exist two constants β and C, only depending on d, Λ, λ_0 and M, such that $|u(x) - u(y)| \le C|x - y|^\beta(1 + \sup_{\overline{B}_2}(|u|))$ for any $x, y \in \overline{B}_1$.

To the best of our knowledge, this result is new: divergence form equations excepted, all the estimates we know for the viscosity solutions of possibly degenerate fully non-linear elliptic PDEs take into account the moduli of continuity of the coefficients. (See, among others, Barles and Da Lio [2], Ishii and Lions [12], Jakobsen and Karlsen [13] and Katsoulakis [15].) Here, the final Hölder bound doesn't depend on the regularity of A and f (despite A and f are assumed to be continuous). Obviously, Theorem 1.1 applies to the p-Laplace operator, $p > 2$, which expands in a non-divergence form as $\Delta_p(u) = |Du|^{p-2}\mathrm{Tr}[(I_d + (p-2)|Du|^{-2}DuDu^*)D^2u]$. ($I_d$ is the identity matrix of size d.) Indeed, Δ_p fulfills (H1) with $\lambda(z) = |z|^{p-2}$ and $\Lambda = p - 1$.

[1] The coefficient 1/2 related to the growth of f in (H1) is purely cosmetic: when λ is $[0, 1]$-valued, $|f(x, y, z)| \le \Lambda(1 + |z|)$.

Before we discuss the strategy of the proof, we say a little bit more about the equation itself. We first emphasize that (1.1) fulfills a maximum principle under Assumption (H1): the supremum norm of the solution u on $\overline{B}_2$ may be bounded in terms of the parameters Λ, λ_0 and M and of the supremum of $|u|$ on the boundary of $\overline{B}_2$ (see e.g. Serrin and Pucci [24, Thm 2.3.2], the proof being easily adapted to viscosity solutions). On the contrary, (H1) is not sufficient to guarantee a strong Harnack inequality of the form $\sup_{B_1}(u) \leq C \inf_{B_1}(u)$ when u is non-negative and f is zero: think of $d = 1$, $u(x) = 1 + \cos(\pi x)$ and $\lambda(z) = 0$ for $|z| \leq \pi$. Concerning the assumptions, we emphasize that the continuity property of the coefficients A, f and λ in the statement could be relaxed: this would demand an additional effort which seems useless here. On the opposite, the optimality of (H1) is to be understood: is it possible to require (H1) only for $\xi = z$ as for divergence form equations? The possible extension of the result to fully non-linear equations of the form $F(x, u, Du, D^2u) = 0$ on the model of the works of Caffarelli [4] and Caffarelli and Cabré [5] on Krylov and Safonov estimates for uniformly elliptic non-linear PDEs is also to be considered. Finally, we emphasize that we haven't been able to adapt the approach to parabolic equations: the problem is to fit the time and space scales properly in the method developed below. This seems to be quite challenging: when the equation degenerates, the natural diffusive scaling between time and space breaks down since the solution locally generates its own scaling according to the values of the diffusion coefficient. Similar difficulties occur for parabolic equations of divergence form: we refer to the series of papers by DiBenedetto, Urbano and Vespri mentioned in their common work [10] for an overview of the method used in that case.

Strategy. As in the original work of Krylov and Safonov for linear equations, the strategy of the proof relies on a probabilistic interpretation of the quasi-linear PDE. Indeed, when A and f are independent of y and z, i.e. when the equation is linear, the original proof consists in introducing a diffusion process X, solution to the Stochastic Differential Equation (SDE for short)

$$dX_t = \sigma(X_t)dW_t, \quad t \geq 0,$$

where W is a d-dimensional Wiener process and σ a continuous version of the square root of the matricial mapping $2A$. (In the linear framework, (H1) ensures that A is elliptic so that the above equation is weakly solvable, see Stroock and Varadhan [26].)

The basic idea of Krylov and Safonov follows from a key observation in the theory of diffusion processes: the generator of a diffusion process enjoys some smoothing property if the paths of the corresponding process sufficiently visit the surrounding space with a non trivial probability. The argument may be understood as follows in the simple case when f vanishes and u is smooth: in such a framework, $(u(X_t))_{t\geq 0}$ is a martingale. In particular, $u(x)$ may be expressed as the expectation $\mathbb{E}[u(X^x_\tau)]$ for any well-controlled stopping time τ. (Here, the exponent x indicates the initial position of the diffusion process.) As a consequence, $u(x)$ may be understood as a mean over the values of u in a neighborhood of x: since X visits the surrounding space around x, almost all the values of u in the neighborhood of x have a role in the computation of the expectation. Obviously, the same is true for any point y very close to x: both $u(x)$ and $u(y)$ may be expressed as expectations over the values of u in the neighborhood of x (and thus of y). Therefore, $u(x)$ and $u(y)$ are close if the values of u in both expectations are averaged with quite similar weights: this is the case if the way the process visits the surrounding space has some uniformity with respect to the starting point. The method also applies when the source term f is non-zero. In this case, the probabilistic representation formula has the form

$$u(x) = \mathbb{E}\Big[u(X^x_\tau) + \int_0^\tau f(X^x_s)ds\Big]. \tag{1.2}$$

In the specific Krylov and Safonov theory, the point is to bound from below the probability that the diffusion process X hits a Borel subset of non-zero Lebesgue measure included in B_2 (or in a smaller ball) before leaving it. Obviously, the ellipticity property plays a crucial role: indeed, if the diffusion matrix A degenerates on an open subset of B_2, there is no chance for X to move inside along the directions of degeneracy.

To handle the possible degeneracies in the non-linear framework, the idea we here develop is the following. When A and u are smooth, we can define X similarly as above by setting:

$$dX_t = \sigma(X_t, u(X_t), Du(X_t))dW_t, \quad t \geq 0,$$

$(x, y, z) \mapsto \sigma(x, y, z)$ being a smooth version of the square root of $2A$. (We do not discuss the existence of this smooth version at this stage of the paper.) When $|Du|$ is large, the assumption (H1) turns into an ellipticity condition, so that the Krylov and Safonov theory applies. Anyhow, because of the possible degeneracies of $A(x, y, z)$ for $|z|$ small, the process may not move inside the part of the space where the gradient $|Du|$ is small. In what follows, we specifically show that we can force the stochastic system on the areas of degeneracy by an additional drift to push it towards the desired Borel subset. Precisely, we will show the following

Theorem 1.2 *Let $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ be a Lipschitz continuous mapping such that*

$$\forall x, \xi \in \mathbb{R}^d, \quad \Lambda^{-1}\hat{\lambda}(x)|\xi|^2 \leq \langle \xi, a(x)\xi\rangle \leq \Lambda\hat{\lambda}(x)|\xi|^2, \quad a(x) = \sigma\sigma^*(x), \tag{H2}$$

for some $\Lambda \geq 1$ and some mapping $\hat{\lambda} : \mathbb{R}^d \to [0, 1]$. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ also denote a filtered probability space satisfying the usual conditions endowed *with an* $(\mathcal{F}_t)_{t\geq 0}$ *Brownian motion* $(W_t)_{t\geq 0}$, α *be a positive real and* Q_1 *be some hypercube of* $\mathbb{R}^d$ *of radius 1. (For our purpose, we prefer hypercubes to Euclidean balls.)*

Then, for any μ *in* $(0,1)$, *there exist some positive constants* $\varepsilon(\mu)$, $R(\mu)$ *and* $(\Gamma_p(\mu))_{1\leq p<2}$, *only depending on* d, α, Λ *and* μ *(and not on* $\hat{\lambda}$ *in (H2) and on* Q_1*), such that, for any* ρ *in* $(0,1)$ *and* x_0 *in* $Q_{\rho/8}$ *(hypercube of* $\mathbb{R}^d$ *of same center as* Q_1 *but of radius* $\rho/8$*), we can find an integrable* d*-dimensional* $(\mathcal{F}_t)_{t\geq 0}$ *progressively measurable process* $(b_t)_{t\geq 0}$ *such that both* $(b_t)_{t\geq 0}$ *and the process* X, *solution to the SDE*

$$X_t = x_0 + \int_0^t b_s ds + \int_0^t \sigma(X_s)dB_s, \quad t \geq 0,$$

fulfill
$$\begin{cases} \forall t \geq 0, \quad \hat{\lambda}(X_t) \geq \alpha \Rightarrow b_t = 0, \\ \forall p \in [1,2), \quad \mathbb{E}\displaystyle\int_0^{+\infty} |b_t|^p dt \leq \Gamma_p(\mu)\rho^{p-2}, \end{cases}$$

and, for any Borel subset $V \subset Q_\rho$ *(hypercube of same center as* Q_1 *but of radius* ρ*)*

$$|Q_\rho \setminus V| < \mu|Q_\rho| \Rightarrow \mathbb{P}\{T_V < (R(\mu)\rho^2) \wedge S_{Q_\rho}\} \geq \varepsilon(\mu),$$

T_V *being the first hitting time of* V *and* S_{Q_ρ} *the first exit time from* Q_ρ *by* X. *(*$|\cdot|$ *here stands for the Lebesgue measure.)*

Comments. Theorem 1.2 may be a bit complicated to understand at first sight. We first emphasize that (H2) is not a strict ellipticity assumption since $\hat{\lambda}$ may vanish: what is important is that, at any x, all the eigenvalues of $a(x)$ behave in the same way. In the specific case when the matrix a is uniformly non-degenerate, the mapping $\hat{\lambda}$ in the statement may be assumed to be equal to 1 without loss of generality. Choosing $\alpha = 1$ in the statement, we observe that the drift $(b_t)_{t\geq 0}$ given by Theorem 1.2 is then always zero, so that X is simply the solution of the SDE $dX_t = \sigma(X_t)dW_t$, $t \geq 0$, with $x_0 \in Q_{\rho/8}$ as initial condition. Theorem 1.2 then says that the probability of hitting a Borel subset V of Q_ρ before leaving the ball Q_ρ is bounded from below by a constant only depending on d and Λ and on the proportion of V in Q_ρ: this exactly fits the original Krylov and Safonov result. (See [17].) When σ degenerates, Theorem 1.2 says that we can force the stochastic system by an additional drift to preserve the Krylov and Safonov result. The connection with Theorem 1.1 may be understood as follows: when u is a strong solution of the PDE (1.1), we choose $a(x)$ in the statement of Theorem 1.2 as $2A(x,u(x),Du(x))$: under (H1), it satisfies $(2\Lambda)^{-1}\lambda(Du(x)) \leq \langle \xi, a(x)\xi\rangle \leq 2\Lambda\lambda(Du(x))$, with λ as in the statement of Theorem 1.1. The term $\lambda(Du(x))$ then plays the role of $\hat{\lambda}(x)$ in Theorem 1.2 (forget for the moment the fact that $\hat{\lambda}$ has to be $[0,1]$-valued): by choosing α in the statement of Theorem 1.2 equal to λ_0 given by (H1) in

Theorem 1.1, we deduce that $|Du(X_t)| \geq M \Rightarrow \lambda(Du(X_t)) \geq \lambda_0 \Rightarrow \hat{\lambda}(X_t) \geq \alpha \Rightarrow b_t = 0$. In other words, the resulting drift $(b_t)_{t\geq 0}$ just acts when the gradient is small, i.e. is bounded by M.

Of course, adding a non-zero drift in the SDE satisfied by $(X_t)_{t\geq 0}$ breaks down the natural connection with the PDE (1.1). Actually, still assuming that u is smooth, a simple application of the Itô formula shows that, for $(X_t)_{t\geq 0}$ as in the previous paragraph, (1.2) becomes

$$u(x_0) = \mathbb{E}\Big[u(X_\tau) + \int_0^\tau \Big(f(X_s) - \langle b_s, Du(X_s)\rangle\Big) ds\Big]. \tag{1.3}$$

Again, τ is a well-controlled stopping time and x_0 is some initial condition as in the statement of Theorem 1.2. Here is the main issue: the best bound we have on $(Du(X_t))_{t\geq 0}$ in such problems holds in $L^2(\Omega, L^2([0,\tau], \mathbb{R}^d))$ (think of the Itô isometry or refer to the more general results on Backward SDEs in which such controls are frequently used, see Pardoux [23] or Delarue [8]); moreover, by Theorem 1.2, the drift $(b_t)_{t\geq 0}$ is just $L^p(\Omega, L^p([0,\tau], \mathbb{R}^d))$ integrable for $1 \leq p < 2$. Therefore, without any additional information on $(b_t)_{t\geq 0}$, there is no hope to give a sense to (1.3). Anyhow, because of its specific construction, $(b_t)_{t\geq 0}$ vanishes for $|Du(X_t)| \geq M$, so that $(|\langle b_t, Du(X_t)\rangle|)_{t\geq 0}$ is always bounded by $(M|b_t|)_{t\geq 0}$. The $L^p(\Omega, L^p([0,\tau], \mathbb{R}^d))$ controls, $1 \leq p < 2$, we have on $(b_t)_{t\geq 0}$ are then sufficient to see (1.3) as a variation of (1.2). It is then possible to derive the estimates for u as in the original paper by Krylov and Safonov [2].

Organization of the Paper. In Section 2, we show how to deduce Theorem 1.1 from Theorem 1.2. In Section 3, we prove Theorem 1.2 when the proportion of V inside Q_ρ is large enough: we call this step "attainability of large sets". This is the core of the proof. It is the equivalent of the first step in the Krylov and Safonov proof: large sets are there shown to be attainable with a non-zero probability by an application of the Krylov inequality. We then complete the proof of Theorem 1.2 in Section 4 by proving that small sets are also attainable: as in the original proof, we first prove that small balls are attainable. Combining the attainability of small balls and the attainability of large sets, we complete the proof.

2 Application of Theorem 1.2 to Degenerate Elliptic Equations

We first show how to derive Theorem 1.1 from Theorem 1.2. Dividing (1.1) by $1+\lambda(Du(x))$, we can assume λ to be $[0,1]$-valued in the whole demonstration. (Obviously, this doesn't change the values of Λ and M and just turns λ_0 into

[2] The form (1.3) explains why the strong Harnack inequality fails. The term $(\langle b_s, Du(X_s)\rangle)_{0\leq s\leq\tau}$ behaves as a non-trivial source term.

$\lambda_0/(1+\lambda_0)$.) For notational simplicity, we also restrict the proof to the case when A and f are independent of y: the argument is completely similar when A and f do depend on y. We thus write (1.1) under the form

$$-\mathrm{Tr}(A(x, Du(x))D^2u(x)) + f(x, Du(x)) = 0, \quad x \in B_2. \tag{2.1}$$

Compared with the original argument given by Krylov and Safonov, the main difference in the application of the probabilistic estimate follows from the interpretation of the underlying PDE. In the paper by Krylov and Safonov, the PDE is understood in the strong sense, i.e. the solution u is assumed to be in $\mathcal{C}(\overline{B}_2) \cap W^{2,d}_{\mathrm{loc}}(B_2)$. We here consider the equation (2.1) in the viscosity sense. The idea to recover the strong framework is classical in the theory of viscosity solutions and consists of a regularization by infimum and supremum convolutions. We refer to the articles by Lasry and Lions [21], Crandall et al. [6] and Jensen [14] for the original ideas. Basically, the infimum and supremum convolutions permit both to regularize a given solution of (2.1) and to keep the original structure of the PDE. Following [21], we thus define for any bounded and uniformly continuous function w on the whole $\mathbb{R}^d$, $\epsilon > 0$ and $x \in \mathbb{R}^d$

$$w^\epsilon(x) = \sup_{y \in \mathbb{R}^d} \Big[w(y) - \frac{1}{2\epsilon}|x-y|^2 \Big], \quad w_\epsilon(x) = \inf_{y \in \mathbb{R}^d} \Big[w(y) + \frac{1}{2\epsilon}|x-y|^2 \Big].$$

The main result in [21] says that, for any positive δ and ϵ, $(w^{\epsilon+\delta})_\delta$ belongs to $\mathcal{C}^{1,1}(\mathbb{R}^d)$ (i.e. is continuously differentiable on $\mathbb{R}^d$ with Lipschitz continuous derivatives) and uniformly converges towards w as δ and ϵ tend to zero. The point is thus to prove that, when w satisfies a given second order PDE in the viscosity sense, $(w^{\epsilon+\delta})_\delta$ is a viscosity subsolution of some PDE similar to the original one. The proof of the following result is inspired from the paper [6]:

Proposition 2.1 *Let A and f be coefficients independent of y fulfilling (H1) with respect to some $\Lambda \geq 1$, $\lambda : \mathbb{R}^d \to [0,1]$, $\lambda_0 \in (0,1]$ and $M > 0$. Let $u : \overline{B}_2 \to \mathbb{R}$ be also a continuous viscosity solution of the PDE (2.1). Setting $w = (\tilde{u}^{\epsilon+\delta})_\delta$ for $\delta > 0$ and $\epsilon > 0$ and for some arbitrarily chosen bounded and uniformly continuous extension $\tilde{u}$ of u to the whole $\mathbb{R}^d$, there exists $\theta \in (0,1)$ such that, for $\delta = \theta\epsilon$ and for ϵ small enough, w satisfies:*

$$-\mathrm{Tr}(A_\epsilon(x, Dw(x))D^2w(x)) + f(x+\epsilon Dw(x), Dw(x)) \leq 2(\Lambda + M) \text{ a.e. } x \in B_{3/2},$$

where $B_{3/2}$ is the ball of same center as B_2 but of radius $3/2$ and $A_\epsilon : \mathbb{R}^d \times \mathbb{R}^d \to \mathcal{S}_d(\mathbb{R})$ is a smooth function (i.e. $\mathcal{C}^\infty$ with bounded derivatives of any order) satisfying $\inf\{\langle \xi, A_\epsilon(x,z)\xi \rangle,\ x, z, \xi \in \mathbb{R}^d, |\xi| = 1\} > 0$ as well as Assumption (H1) with respect to Λ and to some mapping $\lambda_\epsilon : \mathbb{R}^d \to (0,2]$ such that $\lambda_\epsilon(z) \geq \lambda_0$ for $|z| \geq M+1$.

Proof. By Lasry and Lions [21], there exists some constant $K \geq 0$ (K depends on u but is independent of δ and ϵ) such that $|Dw(x)| \leq K\epsilon^{-1/2}$ for any

$x \in B_{3/2}$. We deduce that for ϵ small enough, $x + \epsilon Dw(x)$ belongs to B_2 for any $x \in B_{3/2}$.

Choose now some $\hat{x} \in B_{3/2}$ at which $D^2 w(\hat{x})$ exists. Assume first that $w(\hat{x}) = \tilde{u}^\epsilon(\hat{x})$. Since $w \geq \tilde{u}^\epsilon$ on the whole $\mathbb{R}^d$ (see [21]), $(Dw(\hat{x}), D^2 w(\hat{x})) \in \mathcal{D}^{2,+}\tilde{u}^\epsilon(\hat{x})$ (see Crandall, Ishii and Lions [7] for the standard definition of $\mathcal{D}^{2,+}$). By Crandall et al. [6, Prop 4.3], we deduce that $(Dw(\hat{x}), D^2 w(\hat{x})) \in \mathcal{D}^{2,+}u(\hat{x} + \epsilon Dw(\hat{x}))$, so that, by (2.1)

$$-\mathrm{Tr}\Big(A(\hat{x} + \epsilon Dw(\hat{x}), Dw(\hat{x}))D^2 w(\hat{x})\Big) + f\Big(\hat{x} + \epsilon Dw(\hat{x}), Dw(\hat{x})\Big) = 0. \quad (2.2)$$

Suppose now that $w(\hat{x}) > \tilde{u}^\epsilon(\hat{x})$. By [6, Prop 4.4], $1/\delta$ is an eigenvalue of $D^2 w(\hat{x})$ (and is in fact the largest one) and, by [6, Prop 4.5], the other eigenvalues are greater than or equal to $-1/\epsilon$. In particular, for any $y \in B_{3/2}$,

$$\begin{aligned} &-\mathrm{Tr}\Big(A(y, Dw(\hat{x}))D^2 w(\hat{x})\Big) + f\Big(y, Dw(\hat{x})\Big) \\ &\quad \leq \lambda(Dw(\hat{x}))\Big(-\Lambda^{-1}\delta^{-1} + (d-1)\Lambda\epsilon^{-1}\Big) + \Lambda(1 + |Dw(\hat{x})|). \end{aligned}$$

We now choose $\delta = \epsilon/(d\Lambda^2)$, so that, for any $y \in B_{3/2}$,

$$-\mathrm{Tr}\Big(A(y, Dw(\hat{x}))D^2 w(\hat{x})\Big) + f\Big(y, Dw(\hat{x})\Big) \leq \Lambda\Big[-\lambda(Dw(\hat{x}))\epsilon^{-1} + 1 + |Dw(\hat{x})|\Big]. \quad (2.3)$$

We know that $\lambda(Dw(\hat{x})) \geq \lambda_0$ for $|Dw(\hat{x})| \geq M$, so that the above right-hand side is less than $\Lambda(-\lambda_0\epsilon^{-1} + 1 + K\epsilon^{-1/2})$ for $|Dw(\hat{x})| \geq M$. Otherwise, it is less than $\Lambda(1 + M)$. Choosing ϵ small enough, we deduce from (2.2) and (2.3) that in any case

$$-\mathrm{Tr}\Big(A(\hat{x} + \epsilon Dw(\hat{x}), Dw(\hat{x}))D^2 w(\hat{x})\Big) + f\Big(\hat{x} + \epsilon Dw(\hat{x}), Dw(\hat{x})\Big) \leq \Lambda(1 + M).$$

We finally smooth the diffusion coefficient using a standard mollifier. We know that the norm of $D^2 w(\hat{x})$ is less that some constant $C(\epsilon) \geq 1$. Then, we can find a smooth matricial function $A_\epsilon : \mathbb{R}^d \times \mathbb{R}^d \to \mathcal{S}_d(\mathbb{R})$ such that

$$\sup_{|x| \leq 2, |z| \leq K\epsilon^{-1/2}} |A_\epsilon(x, z) - A(x, z)| \leq \epsilon/C(\epsilon).$$

It is clear that A_ϵ fulfills (H1) with respect to Λ and to some mapping $\lambda_\epsilon : \mathbb{R}^d \to [0, 1]$ obtained by mollification of the original mapping λ. In particular, for ϵ small enough, $\lambda_\epsilon(z) \geq \lambda_0$ if $|z| \geq M + 1$. Changing A_ϵ into $A_\epsilon + [\epsilon/C(\epsilon)]I_d$, we can assume that the lowest eigenvalue of $A_\epsilon(x, z)$ is greater than or equal to $\epsilon/C(\epsilon)$ for any $(x, z) \in \mathbb{R}^d \times \mathbb{R}^d$. This changes λ_ϵ into $\lambda_\epsilon + \epsilon/C(\epsilon)$: for ϵ small enough, the new λ_ϵ is thus $(0, 2]$-valued. For this final choice of A_ϵ,

$$-\mathrm{Tr}\Big(A_\epsilon(\hat{x} + \epsilon Dw(\hat{x}), Dw(\hat{x}))D^2 w(\hat{x})\Big) + f\Big(\hat{x} + \epsilon Dw(\hat{x}), Dw(\hat{x})\Big) \leq 2\epsilon + \Lambda(1 + M).$$

This completes the proof. □

We are now in position to complete the proof of Theorem 1.1. According to Gilbarg and Trudinger [11, Lem 8.23], it is sufficient to prove

Proposition 2.2 *Under the assumptions of Proposition 2.1, there exist two constants $\gamma \in (0,1)$ and $C \geq 0$, only depending on d, Λ, λ_0 and M, such that for any $\rho \in (0,1)$, for any hypercubes $Q_{\rho/8}$ and Q_ρ of same center and of radii $\rho/8$ and ρ, with $Q_{\rho/8} \subset Q_\rho \subset B_{3/2}$ ($B_{3/2}$ being the ball of same center as B_2 in the statement of Theorem 1.1 but of radius 3/2)*

$$\operatorname*{osc}_{Q_{\rho/8}}(u) \leq \gamma \operatorname*{osc}_{Q_\rho}(u) + C\rho(1 + \sup_{Q_\rho}(|u|)) \quad (\operatorname*{osc}_{Q_r}(u) = \sup_{Q_r}(u) - \inf_{Q_r}(u)).$$

Proof. We set $m_- = \inf_{Q_\rho}(u)$ and $m_+ = \sup_{Q_\rho}(u)$. Changing u into $-u$ if necessary, we can assume that $|\{x \in Q_\rho : u(x) \leq (m_+ + m_-)/2\}| \geq (1/2)|Q_\rho|$. We also consider w given by Proposition 2.1, with ϵ and δ as in the statement of the proposition. Changing λ_ϵ into $\lambda_\epsilon/4$, Λ into 4Λ and λ_0 into $\lambda_0/4$, we can assume that λ_ϵ is $(0,1/2]$ valued. Then, we can apply Theorem 1.2 to $a_\epsilon(x) = 2A_\epsilon(x, Dw(x))$. Indeed, since A_ϵ is smooth and uniformly non-degenerate and Dw is Lipschitz continuous, the symmetric square root σ of a_ϵ is also Lipschitz continuous: (H2) in Theorem 1.2 is then easily checked with $\hat{\lambda}(x) = 2\lambda_\epsilon(Dw(x))$. Obviously, the hypercubes to which the theorem is applied are $Q_{\rho/8}$ and Q_ρ and the initial condition x_0 is some arbitrary point in $Q_{\rho/8}$. Moreover, the parameters α and μ are respectively chosen equal to λ_0 and to $1/2$. The resulting processes are denoted by $(b_t)_{t\geq 0}$ and $(X_t)_{t\geq 0}$ and the constants $R(1/2)$ and $\varepsilon(1/2)$ are denoted by R_0 and ε_0.

We then consider $V = \{x \in Q_\rho : u(x) \leq (m_+ + m_-)/2\}$. Using the notations of Theorem 1.2, we also define τ as the sopping time $T_V \wedge (R_0\rho^2) \wedge S_{Q_\rho}$. We wish to apply Itô's formula to $(w(X_t))_{t\geq 0}$. The point is that w is not in $\mathcal{C}^2(\mathbb{R}^d)$ but in $\mathcal{C}^{1,1}(\mathbb{R}^d)$. Since the diffusion matrix of X is uniformly elliptic, we have in mind to apply the Itô-Krylov formula that holds for functions with Sobolev derivatives (see [16, Sec 2.10]). There is then another problem: it requires the drift $(b_t)_{t\geq 0}$ to be bounded. We thus define, for any $n \geq 1$, the Itô process

$$X_t^n = x_0 + \int_0^t \sigma(X_s)dW_s + \int_0^t b_s \mathbf{1}_{\{|b_s|\leq n\}} ds.$$

Since $(b_t)_{t\geq 0}$ belongs to $L^1(\Omega \times \mathbb{R}_+)$, it is clear that $\mathbb{E}[\sup_{t\geq 0}|X_t^n - X_t|]$ tends to 0 with n. Expanding $(w(X_t^n))_{t\geq 0}$ and taking the expectation (τ is bounded):

$$\begin{aligned} w(x_0) &= \mathbb{E}[w(X_\tau^n)] \\ &\quad - \mathbb{E}\int_0^\tau \Big[(1/2)\mathrm{Tr}(a_\epsilon(X_s)D^2 w(X_s^n)) + \langle b_s, Dw(X_s^n)\rangle \mathbf{1}_{\{|b_s|\leq n\}}\Big] ds. \end{aligned}$$

Since w is a subsolution of the regularized version of (1.1), see Proposition 2.1, $-(1/2)\mathrm{Tr}(a_\epsilon(X_s^n)D^2 w(X_s^n)) \leq -f(X_s^n + \epsilon Dw(X_s^n), Dw(X_s^n)) + 2(\Lambda + M)$.

Hence

$$w(x_0) \le \mathbb{E}[w(X^n_\tau)] + \mathbb{E}\int_0^\tau \Big[3\Lambda + 2M + \Lambda|Dw(X^n_s)| \\ - \langle b_s, Dw(X^n_s)\rangle \mathbf{1}_{\{|b_s|\le n\}} + K_\epsilon |X^n_s - X_s|\Big] ds, \tag{2.4}$$

where K_ϵ is a constant depending on the Lipschitz constant of a_ϵ and on the bound of $D^2 w$. It is then plain to let n tend to $+\infty$ in (2.4). The only problem is to get an estimate for the integral of $Dw(X^n_s)$. Setting $v = w - \inf_{B_2}(w)$, it is easily checked that v^2 satisfies $-\mathrm{Tr}(A_\epsilon(y, Dw(y))D^2(v^2)(y)) + 2\langle A_\varepsilon(y, Dw(y))Dw(y), Dw(y)\rangle + 2v(y)f(y + \epsilon Dw(y), Dw(y)) \le 4v(y)(\Lambda + M)$ for a.e. $y \in B_{3/2}$. Repeating the proof of (2.4) for v^2 and letting n tend to $+\infty$, we obtain

$$\begin{aligned} &v^2(x_0) + \mathbb{E}\int_0^\tau \langle a_\varepsilon(X_s)Dw(X_s), Dw(X_s)\rangle ds \\ &\le \mathbb{E}[v^2(X_\tau)] + 2\mathbb{E}\int_0^\tau v(X_s)\Big[3\Lambda + 2M + \Lambda|Dw(X_s)| - \langle b_s, Dw(X_s)\rangle\Big] ds. \end{aligned}$$

Recall that $\langle a_\epsilon(X_s)Dw(X_s), Dw(X_s)\rangle \ge 2\Lambda^{-1}\lambda_0|Dw(X_s)|^2$ if $|Dw(X_s)| \ge M + 1$. Moreover, $|Dw(X_s)| \ge M + 1 \Rightarrow \hat{\lambda}(X_s) \ge \lambda_0 = \alpha \Rightarrow b_s = 0$. It is plain to deduce that there exists a constant C, only depending on Λ, λ_0 and M, such that:

$$\mathbb{E}\int_0^\tau |Dw(X_s)|^2 ds \le \mathbb{E}[v^2(X_\tau)] + C\mathbb{E}\int_0^\tau \Big(1 + v^2(X_s) + v(X_s)|b_s|\Big) ds.$$

By the bounds we have on τ ($\tau \le R_0\rho^2$) and $(|b_t|)_{t\ge 0}$ (see Theorem 1.2), we can bound the right-hand side by $C(1 + \sup_{Q_\rho}(v^2))$ and thus by $C(1 + \sup_{Q_\rho}(w^2))$ (up to a new value of C possibly depending on d). Plugging this bound in (2.4) (with $n \to +\infty$ and with the same trick as above to bound $\langle b_s, Dw(X_s)\rangle$), we obtain (the value of C may vary from line to line)

$$\begin{aligned} w(x_0) &\le \mathbb{E}[w(X_\tau)] + C\mathbb{E}\int_0^\tau \Big[1 + |Dw(X_s)|\Big] ds \\ &\le \mathbb{E}[w(X_\tau)] + C\mathbb{E}(\tau) + \mathbb{E}\Big[\tau^{1/2}\Big(\int_0^\tau |Dw(X_s)|^2 ds\Big)^{1/2}\Big] ds \\ &\le \mathbb{E}[w(X_\tau)] + C\rho(1 + \sup_{Q_\rho}(|w|)) \end{aligned}$$

since $\tau \le R_0\rho^2$. We finally let ϵ tend to 0: w tends to u, uniformly on $\overline{B}_{3/2}$. Hence,

$$u(x_0) \le \mathbb{E}[u(X_\tau)] + C\rho(1 + \sup_{Q_\rho}(|u|)).$$

The result is now clear: with probability greater than or equal to ε_0, X_τ is in V so that $u(X_\tau) \le (m_+ + m_-)/2$; when X_τ is not in V, $u(X_\tau) \le m_+$. Thus, $u(x_0) \le \varepsilon_0(m_+ + m_-)/2 + (1 - \varepsilon_0)m_+ + C\rho(1 + \sup_{Q_\rho}(|u|))$. Finally, $u(x_0) - m_- \le (1 - \varepsilon_0/2)(m_+ - m_-) + C\rho(1 + \sup_{Q_\rho}(|u|))$. This is true for any $x_0 \in Q_{\rho/8}$ so that $\mathrm{osc}_{Q_{\rho/8}}(u) \le (1 - \varepsilon_0/2)\mathrm{osc}_{Q_\rho}(u) + C\rho(1 + \sup_{Q_\rho}(|u|))$. $\square$

3 Proof of Theorem 1.2: Attainability of Large Sets

We now prove Theorem 1.2 when the proportion of V is large enough. The strategy is the following. If there is enough noise in the system, then the probability of hitting a given Borel subset is bounded away from zero: this is the standard Krylov and Safonov theory. If the noise is too small, we build a drift b to push the process towards the desired area.

The following notations are useful: $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is a filtered probability space enjoying the usual conditions endowed with an $(\mathcal{F}_t)_{t\geq 0}$ Brownian motion $(W_t)_{t\geq 0}$. For any stopping time T, $\mathbb{E}_T$ stands for $\mathbb{E}[\cdot|\mathcal{F}_T]$. (In particular, $\mathbb{E}_0$ means $\mathbb{E}[\cdot|\mathcal{F}_0]$.) For a square integrable $\mathcal{F}_0$-measurable random variable $\xi : \Omega \to \mathbb{R}^d$ and for two $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$-valued progressively-measurable processes $(b_t)_{t\geq 0}$ and $(\sigma_t)_{t\geq 0}$, $(b_t)_{t\geq 0}$ and $(\sigma_t)_{t\geq 0}$ being square integrable, $\mathcal{I}_\xi(b,\sigma)$ denotes the Itô process: $X_t = \xi + \int_0^t b_s ds + \int_0^t \sigma_s dW_s$. When $(b_t)_{t\geq 0}$ and/or $(\sigma_t)_{t\geq 0}$ also depend in a Lipschitz way on a spatial argument in $\mathbb{R}^d$, i.e. $b : \Omega \times \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and/or $\sigma : \Omega \times \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$, $\mathcal{S}_\xi(b,\sigma)$ denotes the solution of the SDE : $X_t = \xi + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s$. Finally, for $z \in \mathbb{R}^d$, $\|z\|$ denotes the ℓ^∞ norm of $\mathbb{R}^d$, i.e. $\|z\| = \sup_{i\in\{1,\dots,d\}} |z_i|$, and, for $\rho > 0$, $Q(z, \rho)$ is the hypercube of center z and radius ρ: $|Q_1| = 2^d$ is the volume of $Q(0, 1)$. (Remind by the way that $\|\cdot\| \leq |\cdot| \leq d^{1/2}\|\cdot\|$, $|\cdot|$ being the Euclidean norm.)

3.1 Noisy Systems

We first provide a very simple rule to determine whether the noise inside the system is sufficient to attain a Borel subset of large measure.

Proposition 3.1 *Let $(b_t)_{t\geq 0}$ and $(\sigma_t)_{t\geq 0}$ be two progressively-measurable processes with values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$ such that $\|b_t\| \leq \rho^{-1}$, $\Lambda^{-1}\lambda_t I_d \leq a_t \leq \Lambda\lambda_t I_d$, $a_t = \sigma_t\sigma_t^*$, $t \geq 0$, for some constants $\rho > 0$ and $\Lambda \geq 1$ and some progressively measurable process $(\lambda_t)_{t\geq 0}$ with values in $[0, 1]$. Let $(X_t)_{t\geq 0}$ also denote the Itô process $\mathcal{I}_{X_0}(b, \sigma)$ for a square-integrable $\mathcal{F}_0$-measurable random variable X_0.*

Then, for every $\eta \in (0, 1)$, there exist two positive constants $\mu(\eta)$ and $\varepsilon(\eta)$, $\varepsilon(\eta) \in (0, 1)$, only depending on d, η and Λ (and not on ρ), such that, for any hypercube $Q_{3\rho}$ of radius 3ρ and any Borel subset $V \subset Q_{3\rho}$ satisfying $|Q_{3\rho} \setminus V| \leq \mu(\eta)\rho^d$, $\mathbb{P}_0\{T_V < \rho^2 \wedge S_{Q_{3\rho}}\} \geq \varepsilon(\eta)$ a.e. on the event $\{X_0 \in Q_\rho, \mathbb{E}_0[\int_0^{\rho^2} \lambda_s ds] \geq \eta\rho^2\}$. ($T_V$ stands for the first hitting time of V and $S_{Q_{3\rho}}$ for the first exit time from $Q_{3\rho}$ by X. Q_ρ is the hypercube of same center as $Q_{3\rho}$ but of radius ρ.)

We establish a first version of Proposition 3.1:

Lemma 3.2 *Keep the notations and assumptions of Proposition 3.1 but allow* $(\|b_t\|)_{t\geq 0}$ *to be bounded by* $3\rho^{-1}$ *instead of* ρ^{-1}*. Then, for all* $R \in (0,1)$, $\eta \in (0,1)$*, there exist two positive constants* $\mu(\eta)$ *and* $\varepsilon(\eta)$*, only depending on d,* η *and* Λ *(and not on* ρ *and* R*), such that, for any Borel subset* $V \subset Q_\rho$ *satisfying* $|Q_\rho \setminus V| \leq \mu(\eta)$, $\mathbb{P}_0\{T_V < (\rho^2 R) \wedge S_{Q_\rho}\} \geq \varepsilon(\eta)$ *a.e. on the event* $\{\mathbb{E}_0[\int_0^{(\rho^2 R)\wedge S_{Q_\rho}} \lambda_s ds] \geq \eta\rho^2\}$*. (Obviously,* S_{Q_ρ} *is the first exit time from* Q_ρ *by* X*. We also notice that* $X_0 \in Q_\rho$ *on the above event.)*

We clearly see the difference between the two statements. In Lemma 3.2, the noise has to be evaluated before the exit time from the hypercube Q_ρ. In Proposition 3.1, the exit phenomenon is forgotten. We now prove Lemma 3.2.

Proof (Lemma 3.2). By a scaling argument, we can assume that $\rho = 1$. Indeed, it is plain to establish the result for $(X_t)_{t\geq 0}$ once it is proven for the process $(\hat{X}_t = \rho^{-1} X_{\rho^2 t})_{t\geq 0}$. Clearly, $(\hat{X}_t)_{t\geq 0}$ satisfies $\mathcal{I}_{\rho^{-1}X_0}((\rho b_{\rho^2 t})_{t\geq 0}, (\sigma_{\rho^2 t})_{t\geq 0})$ and fulfills the assumptions of Lemma 3.2 with $\rho = 1$. In the whole proof, we put ourselves on the event $\{\mathbb{E}_0[\int_0^{R\wedge S_{Q_1}} \lambda_t dt] \geq \eta\}$. Since λ is $[0,1]$-valued and $R \leq 1$, the noise between 0 and $R \wedge S_{Q_1}$ is away from zero with a non-zero probability:

$$\eta \leq \mathbb{E}_0\Big[\int_0^{R\wedge S_{Q_1}} \lambda_t dt\Big] \leq \eta^2 \mathbb{P}_0\Big\{\int_0^{R\wedge S_{Q_1}} \lambda_t dt \leq \eta^2\Big\} + \mathbb{P}_0\Big\{\int_0^{R\wedge S_{Q_1}} \lambda_t dt > \eta^2\Big\}$$
$$= \eta^2 + (1-\eta^2)\mathbb{P}_0\Big\{\int_0^{R\wedge S_{Q_1}} \lambda_t dt > \eta^2\Big\}.$$

We deduce $\mathbb{P}_0\Big\{\int_0^{R\wedge S_{Q_1}} \lambda_t dt > \eta^2\Big\} \geq \eta/(1+\eta)$. We then apply [16, Thm 2, Sec 2, Chp 2] with $F(c,a) = c$ and $c_t = 3$ for all $t \geq 0$. Almost surely,

$$\mathbb{E}_0\left[\int_0^{S_{Q_1}} \exp(-3t)\det{}^{1/d}(a_t)\mathbf{1}_{Q_1\setminus V}(X_t)dt\right] \leq C|Q_1 \setminus V|^{1/d},$$

for some constant C only depending on d (and which may vary from line to line). Allowing C to depend on Λ, we can write (recall that $R \leq 1$ and $a_t \geq \Lambda^{-1}\lambda_t I_d$) $\mathbb{E}_0[\int_0^{R\wedge S_{Q_1}} \lambda_t \mathbf{1}_{Q_1\setminus V}(X_t)dt\Big] \leq C|Q_1 \setminus V|^{1/d}$, so that

$$\eta^2 \mathbb{P}_0\Big\{T_V \geq R \wedge S_{Q_1}, \int_0^{R\wedge S_{Q_1}} \lambda_t dt > \eta^2\Big\} \leq C|Q_1 \setminus V|^{1/d}. \tag{3.1}$$

Having in mind the inequality $\mathbb{P}(B_1) \leq \mathbb{P}(B_1 \cap B_2) + 1 - \mathbb{P}(B_2)$, $B_1, B_2 \in \mathcal{F}$, we deduce from the bound $\mathbb{P}_0\{\int_0^{R\wedge S_{Q_1}} \lambda_t dt > \eta^2\} \geq \eta/(1+\eta)$ and from (3.1):

$$\mathbb{P}_0\{T_V \geq R \wedge S_{Q_1}\} \leq C\eta^{-2}|Q_1 \setminus V|^{1/d} + 1/(1+\eta).$$

If $|Q_1 \setminus V| \leq [\eta^3/(2C(1+\eta))]^d$, then $\mathbb{P}_0\{T_V < R \wedge S_{Q_1}\} \geq \eta/[2(1+\eta)]$. □

Proof (Proposition 3.1.) We are now in position to complete the proof of Proposition 3.1. We thus put ourselves on the event:

$$X_0 \in Q_\rho, \ \mathbb{E}_0\Big[\int_0^{\rho^2} \lambda_s ds\Big] \geq \eta\rho^2. \tag{3.2}$$

Since $X_0 \in Q_\rho$, it is clear that $S_{Q_{3\rho}} < \rho^2 \Rightarrow \sup_{0 \leq t \leq \rho^2} \| \int_0^{t \wedge S_{Q_{3\rho}}} \sigma_s dW_s\| + \int_0^{\rho^2} \|b_s\| ds \geq 2\rho \Rightarrow \sup_{0 \leq t \leq \rho^2} | \int_0^{t \wedge S_{Q_{3\rho}}} \sigma_s dW_s| \geq \rho$. (Indeed, $\|b_t\| \leq \rho^{-1}$, $t \geq 0$.) By Doob's maximal inequality, $\mathbb{P}_0\{S_{Q_{3\rho}} < \rho^2\} \leq \rho^{-2}\mathbb{E}_0[\int_0^{\rho^2 \wedge S_{Q_{3\rho}}} \mathrm{Tr}[a_s] ds]$. By the specific structure of $(a_t)_{t \geq 0}$, $\mathrm{Tr}(a_t) \leq d\Lambda\lambda_t$ for any $t \geq 0$, so that

$$\mathbb{P}_0\{S_{Q_{3\rho}} < \rho^2\} \leq d\Lambda\rho^{-2}\mathbb{E}_0\Big[\int_0^{\rho^2 \wedge S_{Q_{3\rho}}} \lambda_s ds\Big].$$

In particular, if

$$\mathbb{E}_0\Big[\int_0^{\rho^2 \wedge S_{Q_{3\rho}}} \lambda_s ds\Big] < \eta\rho^2/(2d\Lambda), \tag{3.3}$$

then $\mathbb{P}_0\{S_{Q_{3\rho}} < \rho^2\} < \eta/2$, so that (3.2) together with the bound $\lambda_t \leq 1$, $t \geq 0$, yield

$$\mathbb{E}_0\Big[\int_0^{\rho^2 \wedge S_{Q_{3\rho}}} \lambda_s ds\Big] \geq \mathbb{E}_0\Big[\int_0^{\rho^2} \lambda_s ds\Big] - \rho^2\mathbb{P}_0\{S_{Q_{3\rho}} < \rho^2\} > \eta\rho^2/2.$$

Therefore, (3.3) is impossible. In particular, there exists a constant $c > 1$, only depending on d and Λ, such that

$$\mathbb{E}_0\Big[\int_0^{\rho^2 \wedge S_{Q_{3\rho}}} \lambda_s ds\Big] \geq \eta\rho^2/c = \eta/(9c)(3\rho)^2.$$

We finally apply Lemma 3.2 to the hypercube $Q_{3\rho}$ with $R = 1/9$ (we note that $\|b_t\| \leq \rho^{-1} = 3(3\rho)^{-1}$). For $|Q_{3\rho} \setminus V| \leq \mu(\eta/9c)$, $\mathbb{P}_0\{T_V < \rho^2 \wedge S_{Q_{3\rho}}\} \geq \varepsilon(\eta/(9c))$. $\square$

3.2 Remarkable Points in Large Sets

The point now is to understand what happens when the noise is too small. As already explained, we aim at pushing the process by a well-chosen drift towards the Borel subset V. The question is: towards which part of V do we have to push the process? A possible strategy consists in forcing the process X to go to the neighborhood of some remarkable point x in V, given by

Lemma 3.3 *There exist two universal constants $q_0 > 0$ and $K_0 \geq 0$, only depending on d, such that, for any Borel subset $V \subset Q(0,1)$ satisfying $|Q(0,1) \setminus V| \leq q_0$, there exists $x \in Q(0,1/8) \cap V$ such that, for any $\rho \in (0,3/4)$, $|Q(x,\rho) \setminus V| \leq K_0|Q(0,1) \setminus V|^{1/2}\rho^d$.*

What Lemma 3.3 says is the following: if the proportion of V inside $Q(0,1)$ is large enough, then we can find some point x close to zero such that the proportion of V inside any neighborhood of x is also large. Of course, this result is close to the Lebesgue theorem: for a.e. point $z \in V$, we know that $|Q(z,\rho)|^{-1}|Q(z,\rho)\cap V|$ tends to one as ρ tends to zero, so that the proportion of V inside any small neighborhood of z is large. Lemma 3.3 is in fact a bit stronger: the lower bound for the proportion of V inside a given neighborhood of x doesn't depend on the radius of the neighborhood.

Proof (Lemma 3.3). We admit for the moment the following version (pay attention, in what follows, the sets we consider are small sets whereas they are large sets in the statement of Lemma 3.3):

Lemma 3.4 *There exist three universal constants $q_1 \in (0,1)$, $K_1 \geq 0$ and $\delta \in (0,1/2)$, only depending on d, such that, for any Borel subset $U \subset [0,1]^d$ satisfying $|U| \leq q_1$, there exists $y \in [\delta, 1-\delta]^d \cap U^{\complement}$ such that, for any $\rho > 0$, $|U \cap Q(y,\rho)| \leq K_1|U|^{1/2}\rho^d$.*

We then apply Lemma 3.4 to $U = \{z \in [0,1]^d,\ (1/8)z \in V^{\complement}\}$, i.e. $U = 8([0,1/8]^d \cap V^{\complement})$. Then, $|U| = 8^d|[0,1/8]^d \cap V^{\complement}| \leq 8^d|Q(0,1)\setminus V|$. Therefore, with q_1, K_1 and δ as above, for $|Q(0,1)\setminus V| \leq 8^{-d}q_1$,

$$\exists y \in [\delta, 1-\delta]^d \cap U^{\complement} : \forall \rho > 0,\ |U \cap Q(y,\rho)| \leq K_1|U|^{1/2}\rho^d. \tag{3.4}$$

Set $x = (1/8)y \in [\delta/8, 1/8-\delta/8]^d \cap ((1/8)U^{\complement}) \subset Q(0,1/8)\cap V$. For $\rho < \delta/8$, $|Q(x,\rho)\setminus V| = |Q(x,\rho)\cap V^{\complement}| = 8^{-d}|Q(y,8\rho)\cap(8V^{\complement})| = 8^{-d}|Q(y,8\rho)\cap U|$ since $Q(y,8\rho) \subset [0,1]^d$. Therefore, for $\rho < \delta/8$ and $|Q(0,1)\setminus V| \leq 8^{-d}q_1$, (3.4) yields

$$|Q(x,\rho)\setminus V| \leq 8^{-d}K_1|U|^{1/2}(8\rho)^d \leq 8^{d/2}K_1|Q(0,1)\setminus V|^{1/2}\rho^d. \tag{3.5}$$

Finally, for $\delta/8 \leq \rho < 3/4$, $Q(x,\rho) \subset Q(0,1)$ so that

$$|Q(x,\rho)\setminus V| \leq |Q(0,1)\setminus V| \leq (8/\delta)^d|Q_1|^{1/2}|Q(0,1)\setminus V|^{1/2}\rho^d. \tag{3.6}$$

By (3.5)-(3.6), we set $q_0 = 8^{-d}q_1$ and $K_0 = \max(8^{d/2}K_1, (8/\delta)^d|Q_1|^{1/2})$. □

Proof (Lemma 3.4). We start by a simple lemma (below, U is as in the statement of Lemma 3.4):

Lemma 3.5 *Let p be an integer greater than 3 and E be the square $E = [1/p, 1/p+1/2)^d \subset (0,1)^d$. For any integer $n \geq 1$, we denote by $\mathcal{C}_n$ the collection of hypercubes $\mathcal{R}_n(\ell_1,\dots,\ell_d)$ included in E of the form $\mathcal{R}_n(\ell_1,\dots,\ell_d) = \prod_{i=1}^d [1/p+\ell_i/2^n, 1/p+(\ell_i+1)/2^n)$, $0 \leq \ell_i < 2^{n-1}$, $1 \leq i \leq d$, and we put $M_n(x) = \sum_{B\in\mathcal{C}_n}[(|U\cap B|/|B|)\mathbf{1}_B(x)]$ for any $x \in \mathbb{R}^d$. ($M_n(x)$ is the proportion of U inside the hypercube containing x.) Then, $|F| \geq 1/2^d - |U|^{1/2}$, where $F = \{x \in E : \sup_{n\geq 1} M_n(x) \leq |U|^{1/2}\}$.*

Proof (Lemma 3.5). We endow E with the Borel σ-algebra $\mathcal{G} = \mathcal{B}(E)$ and with the probability measure $\mu = 2^d|\cdot|$. For any $n \geq 1$, we also denote by $\mathcal{G}_n$ the σ-subalgebra of $\mathcal{G}$ generated by the collection $\mathcal{C}_n$. (It is in fact an algebra since its cardinality is finite). It is well seen that the sequence $(\mathcal{G}_n)_{n\geq 1}$ is a filtration and that, on E, the sequence $(M_n)_{n\geq 1}$ coincides with the martingale $(\mathcal{E}(\mathbf{1}_{U\cap E}|\mathcal{G}_n))_{n\geq 1}$, where $\mathcal{E}$ stands for the expectation associated with μ. By Doob's maximal inequality, for every $\varepsilon > 0$, $\mu\{x \in E : \sup_{n\geq 1} M_n(x) > \varepsilon\} \leq \varepsilon^{-1}\mu(U \cap E)$. Choosing $\varepsilon = |U|^{1/2}$, we deduce that $|F| \geq |E| - |U|^{-1/2}|U \cap E|$. $\square$

To complete the proof of Lemma 3.4, we let p vary: we choose $p = p_1, \dots, p_{d+1}$, $(p_i)_{1\leq i\leq d+1}$ being odd integers such that $3 \leq p_1 < p_2 < \dots < p_{d+1}$. (The precise value of p_1 is chosen later.) The resulting quantities E, F, $(M_n)_{n\geq 1}$ and $(\mathcal{R}_n(\ell))_{\ell\in\{0,\dots,2^{n-1}-1\}^d}$ in Lemma 3.5 now depend on i: to indicate the dependence, we write E^i, F^i, $(M_n^i)_{n\geq 1}$ and $(\mathcal{R}_n^i(\ell))_{\ell\in\{0,\dots,2^{n-1}-1\}^d}$, $i \in \{1, \dots, d+1\}$. We also denote by q_1 a real in $(0, 1)$ whose value has to be determined: in what follows, $|U|$ is always less than q_1. By Lemma 3.5, for any $i \in \{1, \dots, d+1\}$, $|F^i| \geq 1/2^d - q_1^{1/2}$.

It is clear that $F^i \subset E^i \subset D = [1/p_{d+1}, 1/2+1/p_1)^d$ for each $i \in \{1, \dots, d+1\}$. Then, for each $i \in \{1, \dots, d+1\}$, $|D \setminus F^i| = (1/2 + 1/p_1 - 1/p_{d+1})^d - |F^i| \leq q_1^{1/2} + c/p_1$, where c is a constant only depending on d. Hence,

$$|\bigcap_{i=1}^{d+1} F^i| \geq |D| - \sum_{i=1}^{d+1} |D \setminus F^i| \geq 1/2^d - (d+1)q_1^{1/2} - (d+1)c/p_1.$$

Finally,

$$\begin{aligned}|\bigcap_{i=1}^{d+1} F^i \cap ([0,1]^d \setminus U)| &= |\bigcap_{i=1}^{d+1} F^i| + |[0,1]^d \setminus U| - |\bigcap_{i=1}^{d+1} F^i \cup ([0,1]^d \setminus U)| \\ &\geq 1/2^d - (d+1)q_1^{1/2} - (d+1)c/p_1 + (1-q_1) - 1 \\ &\geq 1/2^d - (d+2)q_1^{1/2} - (d+1)c/p_1.\end{aligned}$$

Assuming that $(d+2)q_1^{1/2} \leq 1/2^{d+2}$ and $(d+1)c/p_1 \leq 1/2^{d+2}$, we deduce that $\bigcap_{i=1}^{d+1} F^i \cap U^\complement$ is not empty: we choose y inside. Of course, y belongs to each E^i, $i \in \{1, \dots, d+1\}$. In particular, for any $n \geq 1$ and $i \in \{1, \dots, d+1\}$, there exists one and only one hypercube $\mathcal{R}_n^i$ of the form $\mathcal{R}_n^i(\ell)$, $\ell \in \{0, \dots, 2^{n-1}-1\}^d$, such that $y \in \mathcal{R}_n^i$. Since y belongs to F^i for each $i \in \{1, \dots, d+1\}$, we have, for any $n \geq 1$ and $i \in \{1, \dots, d+1\}$, $|U \cap \mathcal{R}_n^i| \leq |U|^{1/2}2^{-dn}$.

We wish to prove that, for each $n \geq 1$, y is in the interior of one of the $(d+1)$ hypercubes $(\mathcal{R}_n^i)_{1\leq i\leq d+1}$. What we say is the following. For $n \geq 1$ and $i \in \{1, \dots, d+1\}$, we have $\mathcal{R}_n^i = \mathcal{R}_n^i(\ell^i)$ for some $\ell^i \in \{0, \dots, 2^{n-1}-1\}^d$. To simplify the notations, we set $k^i = \ell_1^i$, the first coordinate of ℓ^i. We notice that, for $i, j \in \{1, \dots, d+1\}$, $i < j$, $1/p_i + k^i/2^n$ and $1/p_j + k^j/2^n$ are to be different (otherwise, $1/p_i = 1/p_j + h/2^n$ for some integer $h \geq 1$, i.e. $2^n(p_j - p_i) = hp_ip_j$: this is absurd since p_j cannot divide $2^n(p_j - p_i)$). Hence, $|1/p_i + k^i/2^n - 1/p_j - k^j/2^n| = |2^n(p_j - p_i) + (k^i - k^j)p_jp_i|/(2^np_ip_j) \geq 1/(2^np_ip_j) \geq 1/(2^np_{d+1}^2)$.

Now, we can find a permutation $i_1, \dots, i_{d+1}$ of $1, \dots, d+1$ such that $y_1 \geq 1/p_{i_1} + k^{i_1}/2^n > 1/p_{i_2} + k^{i_2}/2^n > \cdots > 1/p_{i_{d+1}} + k^{i_{d+1}}/2^n$ ($y_1, \dots, y_d$ are the d coordinates of y). We denote the permutation $(i_1, \dots, i_{d+1})$ by $(i_{1,1}, \dots, i_{1,d+1})$. Similarly, for each $j \in \{2, \dots, d\}$, we can find a permutation $i_{j,1}, \dots, i_{j,d+1}$ of $1, \dots, d+1$ such that

$$y_j \geq 1/p_{i_{j,1}} + \ell_j^{i_{j,1}}/2^n > 1/p_{i_{j,2}} + \ell_j^{i_{j,2}}/2^n > \cdots > 1/p_{i_{j,d+1}} + \ell_j^{i_{j,d+1}}/2^n.$$

It is clear that $\bigcap_{j=1}^d \{i_{j,2}, \dots, i_{j,d+1}\}$ is not empty (the cardinality of the complementary in $\{1, \dots, d+1\}$ is at most equal to d). We choose s in it. It is also clear that $1/p_{i_{1,1}} + k^{i_{1,1}}/2^n > 1/p_s + k^s/2^n$, so that $1/p_{i_{1,1}} + k^{i_{1,1}}/2^n \geq 1/p_s + k^s/2^n + 1/(2^np_{d+1}^2)$ (because of the minimal distance between two of those reals). Finally, we have $y_1 \geq 1/p_s + k^s/2^n + 1/(2^np_{d+1}^2)$ and, more generally, $y_j \geq 1/p_s + \ell_j^s/2^n + 1/(2^np_{d+1}^2)$, $j \in \{1, \dots, d\}$. We deduce that $\prod_{j=1}^d (y_j - 1/(2^np_{d+1}^2), y_j]$ is included in $\mathcal{R}_n^s$. (Recall that $y \in \mathcal{R}_n^s$.) Hence,

$$\Big|U \cap \prod_{j=1}^d (y_j - 1/(2^np_{d+1}^2), y_j]\Big| \leq |U \cap \mathcal{R}_n^s| \leq |U|^{1/2}2^{-dn}.$$

The same may be shown for each $\prod_{j=1}^d [y_j, y_j + \varepsilon_j/(2^np_{d+1}^2))$, $(\varepsilon_1, \dots, \varepsilon_d) \in \{-1, 1\}^d$ (for possibly different values of s). Finally, $|U \cap Q(y, 1/(2^np_{d+1}^2))| \leq 2^d|U|^{1/2}2^{-dn}$. This is true for all $n \geq 1$.

Choose now $\rho \in (0, 1/(2p_{d+1}^2))$. We can find $n \geq 1$ such that $1/(2^{n+1}p_{d+1}^2) < \rho \leq 1/(2^np_{d+1}^2)$. We then write $|U \cap Q(y, \rho)| \leq |U \cap Q(y, 1/(2^np_{d+1}^2))| \leq 2^d|U|^{1/2}2^{-dn} \leq 4^dp_{d+1}^{2d}|U|^{1/2}\rho^d$. The inequality is still true when $\rho > 1/(2p_{d+1}^2)$: $|U \cap Q(y, \rho)| \leq |U|^{1/2} \leq 4^dp_{d+1}^{2d}|U|^{1/2}\rho^d$. $\square$

3.3 Forcing the System by a Drift

We now prove the main result of this section. Under suitable assumptions on the coefficients, we can build a drift to force the system to hit, with a non-zero probability, a prescribed Borel subset of large measure.

Proposition 3.6 *Let $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ be a Lipschitz continuous mapping such that $a = \sigma\sigma^*$ fulfill (H2) in Theorem 1.2 with respect to some $\Lambda \geq 1$ and $\lambda : \mathbb{R}^d \to [0, 1]$. (No confusion being possible with Theorem 1.1, we forget the "hat" on λ.) Let α be in $(0, +\infty)$. Then, there exist positive constants μ_0, ε_0, R_0 and $(\Gamma_p)_{1\leq p<2}$, only depending on d, α and Λ (and not on λ), such that, for any $\rho \in (0, 1)$, any hypercubes $Q_{\rho/8} \subset Q_\rho \subset \mathbb{R}^d$ of same center and of radii*

$\rho/8$ and ρ and any square integrable $\mathcal{F}_0$-measurable random variable X_0 with values in $\mathbb{R}^d$, there exists an integrable d-dimensional progressively-measurable process $(b_t)_{t\geq 0}$ such that both $(b_t)_{t\geq 0}$ and the process X, equal to $\mathcal{S}_{X_0}((b_t)_{t\geq 0}, \sigma)$, fulfill

$$\begin{cases} \forall t \geq 0, \quad \lambda(X_t) \geq \alpha \Rightarrow b_t = 0, \\ \forall p \in [1,2), \quad \mathbb{E}_0 \displaystyle\int_0^{+\infty} |b_t|^p dt \leq \Gamma_p \rho^{2-p}, \end{cases}$$

and, for any Borel subset $V \subset Q_\rho$ satisfying $|Q_\rho \setminus V| \leq \mu_0 |Q_\rho|$, $\mathbb{P}_0\{T_V < (R_0\rho^2) \wedge S_{Q_\rho}\} \geq \varepsilon_0$ a.e. on the event $\{X_0 \in Q_{\rho/8}\}$. (As usual, T_V is the first hitting time of V and S_{Q_ρ} the first exit time from Q_ρ by X.)

Proof. By a scaling argument, we can assume $\rho = 1$. (See the proof of Lemma 3.2.) In the whole proof, δ denotes a small real (at least less than $1/4$). By $o(1)$, we mean a function of δ that only depends on d, α and Λ (and not on λ) and that tends to zero as δ tends to zero. Changing the values of X_0 if necessary, we can assume that $X_0 \in Q_{1/8}$ a.e. We also assume $|Q_1 \setminus V| \leq q_0$, q_0 being given by Lemma 3.3. Then, we can find a constant $K_0 > 0$, only depending on d, and $x_\infty \in Q_{1/8} \cap V$ such that, for any $r \in (0, 3/4)$,

$$|Q(x_\infty, r) \setminus V| \leq K_0 |Q_1 \setminus V|^{1/2} r^d. \tag{3.7}$$

***Step 1. Construction of** b **and** X.* We first define the following local dynamics. For a finite stopping time T and two $\mathcal{F}_T$-measurable random variables $N : \Omega \to \mathbb{Z}$ and $Y_0 : \Omega \to \mathbb{R}^d$, we define the drift $b_t^{T,Y_0,N} = \delta^{-2N}(x_\infty - Y_0)$, $T \leq t \leq T + \delta^{2n}$. For a smooth function $\psi : \mathbb{R} \to [0,1]$, matching 1 on $(-\infty, \alpha/2]$ and vanishing on $[\alpha, +\infty)$, we then solve the SDE

$$Y_t^{T,Y_0,n} = Y_0 + \int_T^t \psi(\lambda(Y_s^{T,Y_0,N})) b_s^{T,Y_0,N} ds + \int_T^t \sigma(Y_s^{T,Y_0,N}) dW_s, \; T \leq t \leq T + \delta^{2N}.$$

With these notations at hand, we can define $(X_t)_{t\geq 0}$ as follows.

Step 1a. Initialization. We set $T_0 = 0$ (initial time). We know that $X_0 \in Q(x_\infty, 1/4)$ since both x_∞ and X_0 are in $Q_{1/8}$. If $X_0 \neq x_\infty$, there exists a random integer $n_0 \in \mathbb{N}$ such that $X_0 \in Q(x_\infty, \delta^{n_0}) \setminus Q(x_\infty, \delta^{n_0+1})$. We set $T_1 = \delta^{2n_0}$ and $X_t = Y_t^{0,X_0,n_0}$ for $t \in [0, T_1]$. If $X_0 = x_\infty$, we choose $b_t = 0$ for $t \geq 0$ and $(X_t)_{t\geq 0} = \mathcal{S}_{X_0}(0, \sigma)$ and we set $T_{k+1} = +\infty$ and $n_k = +\infty$ for any $k \geq 0$: in this case, the construction is over.

Step 1b. Stop after one step. Assume $n_0 < +\infty$ (otherwise the construction is over). If $X_{T_1} = x_\infty$, we choose $b_t = 0$ for $t \geq T_1$, define $(X_t)_{t\geq T_1}$ as the solution of $X_t = X_{T_1} + \int_{T_1}^t \sigma(X_s) dW_s$, $t \geq T_1$, and set $T_{k+1} = +\infty$ and $n_k = +\infty$ for any $k \geq 1$: the construction is over.

Step 1c. Iteration. Assume $n_0 < +\infty$ and $X_{T_1} \neq x_\infty$. Then, there exists a random integer $n_1 \in \mathbb{Z}$ such that $X_{T_1} \in Q(x_\infty, \delta^{n_1}) \setminus Q(x_\infty, \delta^{n_1+1})$. We then set $T_2 = T_1 + \delta^{2n_1}$ and $X_t = Y_t^{T_1, X_{T_1}, n_1}$ for $t \in [T_1, T_2]$. We then apply Step 1b to X_{T_2}: if $X_{T_2} = x_\infty$, we choose $b_t = 0$ for $t \geq T_2$, define $(X_t)_{t \geq T_2}$ as the solution of $X_t = X_{T_2} + \int_{T_2}^t \sigma(X_s) dW_s$, $t \geq T_2$, and set $T_{k+1} = +\infty$ and $n_k = +\infty$ for any $k \geq 2$. In this case, the construction is over. Otherwise we perform another iteration. And so on...

Step 1d. Notations. Obviously, the random times $(T_k)_{k \geq 0}$ are stopping times. (In short, for any $k \geq 0$, $T_k \leq T_{k+1}$ and T_{k+1} is $\mathcal{F}_{T_k}$ measurable.) We introduce four additional stopping times. We denote by S the first exit time of X from the hypercube $Q(x_\infty, 3/4)$ and we set

$$\tau = \tau_1 \wedge \tau_2, \quad \begin{cases} \tau_1 = \inf\Big\{k \geq 0,\ n_k = +\infty\Big\}, \\ \tau_2 = \inf\Big\{k \geq 0,\ \mathbb{E}_{T_k}\Big[\displaystyle\int_{T_k}^{T_{k+1}} \lambda(X_s) ds\Big] \geq \delta^{2n_k+4}\Big\}. \end{cases} \tag{3.8}$$

(These are discrete stopping times with respect to the filtration $(\mathcal{F}_{T_k})_{k \geq 0}$.) We may explain the role of these stopping times as follows. Using the definition of τ_1, we are first able to summarize the dynamics of $(X_t)_{t \geq 0}$. If $t \in [T_k, T_{k+1})$ with $k \geq \tau_1$, then $dX_t = \sigma(X_t) dW_t$. If $t \in [T_k, T_{k+1})$ with $0 \leq k < \tau_1$, then $T_{k+1} = T_k + \delta^{2n_k}$ and

$$\begin{aligned} dX_t &= \delta^{-2n_k} \psi(\lambda(X_t))(x_\infty - X_{T_k}) dt + \sigma(X_t) dW_t \\ &= (T_{k+1} - T_k)^{-1} \psi(\lambda(X_t))(x_\infty - X_{T_k}) dt + \sigma(X_t) dW_t. \end{aligned} \tag{3.9}$$

The stopping time τ_2 permits to evaluate the noise inside the system and the exit time S to draw a security ball around x_∞: we will show that the process $(X_t)_{t \geq 0}$ hits V before S with a non-zero probability. In this framework, we notice that, for any $k \geq 0$, $n_k \geq 0$ for $T_k < S$. Moreover,

$$\forall 0 \leq k < \tau_1, \quad \begin{cases} n_{k+1} = +\infty \Leftrightarrow X_{T_{k+1}} = x_\infty, \\ n_{k+1} = \ell \Leftrightarrow \delta^{\ell+1} \leq \|X_{T_{k+1}} - x_\infty\| < \delta^\ell, \quad \ell \in \mathbb{Z}. \end{cases} \tag{3.10}$$

Step 1e. Strategy. The strategy now consists in proving that, with a non-zero probability, in a finite time less than S, either there is enough noise in the system or the process X hits x_∞. In both cases, X hits V before leaving the hypercube $Q(x_\infty, 3/4)$ with a non-zero conditional probability. (The word "conditional" means "conditionally to each of both cases".)

The reason why we expect such a behavior may be explained as follows. At a given time $T_k < S$, $0 \leq k < \tau_1$, X_{T_k} is in the hypercube $Q(x_\infty, \delta^{n_k})$, $n_k \geq 0$. If there is enough noise in the system, i.e. $\mathbb{E}_{T_k}\Big[\int_{T_k}^{T_k+\delta^{2n_k}} \lambda(X_s) ds\Big] \geq \delta^{2n_k+4}$, we intend to apply Proposition 3.1 with $\rho = \delta^{n_k}$ and $\eta = \delta^4$: by the specific construction of x_∞, the proportion of V inside the hypercube $Q(x_\infty, 3\delta^{n_k})$ may be chosen as close to 1 as desired by choosing $|Q_1 \setminus V|$ as small as necessary. Therefore, we expect the process X to hit V between times T_k and T_{k+1} with a non-zero conditional probability. (The word "conditional" means "conditionally to $\mathcal{F}_{T_k}$".) If the noise in the system is less than δ^{2n_k+4}, $X_{T_{k+1}}$ is expected to belong to the hypercube $Q(x_\infty, \delta^{n_k+1})$ with a high conditional probability: by the specific construction of X, $X_{T_{k+1}} - x_\infty$ is equal to

$$X_{T_{k+1}} - x_\infty = \delta^{-2n_k} \int_{T_k}^{T_{k+1}} [\psi(\lambda(X_t)) - 1](x_\infty - X_{T_k})dt + \int_{T_k}^{T_{k+1}} \sigma(X_t)dW_t. \tag{3.11}$$

When the noise is small, the conditional expectation of the distance between $X_{T_{k+1}}$ and x_∞ knowing $\mathcal{F}_{T_k}$ is less than (have in mind $\|x_\infty - X_{T_k}\| \leq \delta^{n_k}$, $T_{k+1} - T_k = \delta^{2n_k}$, $\psi(r) = 1$ for $r \leq \alpha/2$ and $\mathrm{Tr}(a(x)) \leq d\Lambda\lambda(x)$)

$$\begin{aligned} &\mathbb{E}_{T_k}\Big[\|X_{T_{k+1}} - x_\infty\|^2\Big] \\ &\leq 2\mathbb{E}_{T_k}\left[\int_{T_k}^{T_{k+1}} \mathbf{1}_{\{\lambda(X_t)>\alpha/2\}}dt\right] + 2\mathbb{E}_{T_k}\left[\int_{T_k}^{T_{k+1}} \mathrm{Tr}[a(X_t)]dt\right] \\ &\leq \Big[4/\alpha + 2d\Lambda\Big]\mathbb{E}_{T_k}\left[\int_{T_k}^{T_{k+1}} \lambda(X_t)dt\right] \leq \Big[4/\alpha + 2d\Lambda\Big]\delta^{2n_k+4}. \end{aligned} \tag{3.12}$$

With a high conditional probability, n_{k+1} is thus expected to be larger than n_k+1. Therefore, if the noise in the system is always small, $(n_k)_{k\geq 0}$ is expected to be at least of linear growth with a non-zero probability. In this case, $X_{T_k} \to x_\infty$ as $k \to +\infty$ and the sequence $(T_k)_{k\geq 0}$ decays at a geometric rate so that $\lim_{k\to+\infty} T_k$ is finite: X hits $x_\infty \in V$ in a finite time. The first step of the proof is thus clear: we have to investigate the growth of the sequence $(n_k)_{k\geq 0}$.

Step 2. Growth of $(n_{k\wedge\tau})_{k\geq 0}$ up to the exit time.

Step 2a. Stochastic comparison. For $0 \leq k < \tau$ and $T_k < S$, we deduce from (3.10) that, for any $\ell \geq 0$, $\{n_{k+1} = n_k - \ell\} = \{\delta^{n_k-\ell+1} \leq \|X_{T_{k+1}} - x_\infty\| < \delta^{n_k-\ell}\}$. By (3.12), on the event $\{\tau > k\} \cap \{T_k < S\}$, for any integer $\ell \geq 0$,

$$\mathbb{P}_{T_k}\{n_{k+1} = n_k - \ell\} \leq C\delta^{-2(n_k-\ell+1)}\delta^{2n_k+4} = C\delta^{2(1+\ell)}, \quad C = 4/\alpha + 2d\Lambda.$$

We thus compare the conditional law of $n_{k+1} - n_k$ knowing $\mathcal{F}_{T_k}$ with the law of some variable ξ_{k+1} with values into $\{\ell \in \mathbb{Z} : \ell \leq 1\}$ such that

$$\mathbb{Q}\{\xi_{k+1} = -\ell\} = C\delta^{2(1+\ell)}, \ \ell \geq 0, \quad \mathbb{Q}\{\xi_{k+1} = 1\} = 1 - C\delta^2/(1-\delta^2),$$

ξ_{k+1} being defined on another probability space $(\Xi, \mathcal{A}, \mathbb{Q})$. (Of course, for δ small enough, $1 - C\delta^2/(1-\delta^2) \geq 0$.)

For $\ell \leq 0$, we have $\mathbb{Q}\{\xi_{k+1} = \ell\} \geq \mathbb{P}_{T_k}\{n_{k+1} - n_k = \ell\}$ on the event $\{\tau > k\} \cap \{T_k < S\}$. This is nothing but saying that the conditional law of $n_{k+1} - n_k$ knowing $\mathcal{F}_{T_k}$ is stochastically less than the law of ξ_{k+1} on the event $\{\tau >$

$k\} \cap \{T_k < S\}$. Indeed, for a non-increasing function $f : \mathbb{R} \to \mathbb{R}$, it is easily checked that $\mathbb{E}^{\mathbb{Q}}[f(\xi_{k+1})] \geq \mathbb{E}_{T_k}[f(n_{k+1} - n_k)]$ on $\{\tau > k\} \cap \{T_k < S\}$.

We now consider a sequence $(\xi_k)_{k\geq 1}$ of I.I.D. random variables with the same law as above. For f non-increasing and non-negative, we have, for any $k \geq 0$,

$$\begin{aligned}\mathbb{E}_0\Big[f(n_{k+1}); \tau > k, T_k < S\Big] &= \mathbb{E}_0\Big[\mathbb{E}_{T_k}[f(n_{k+1})]; \tau > k, T_k < S\Big] \\ &\leq \mathbb{E}_0\Big[\mathbb{E}^{\mathbb{Q}}[f(n_k + \xi_{k+1})]; \tau > k, T_k < S\Big] \\ &\leq \mathbb{E}_0\Big[\mathbb{E}^{\mathbb{Q}}[f(n_k + \xi_{k+1})]; \tau > k-1, T_{k-1} < S\Big].\end{aligned}$$

($T_{-1} = 0$). Noting that the function $y \mapsto \mathbb{E}_{\mathbb{Q}}[f(y + \xi_{k+1})]$ is non-increasing, an induction yields $\mathbb{E}_0[f(n_{k+1}); \tau > k, T_k < S] \leq \mathbb{E}_0[\mathbb{E}^{\mathbb{Q}}[f(n_0 + \xi_1 + \cdots + \xi_{k+1})]] \leq \mathbb{E}^{\mathbb{Q}}[f(\xi_1 + \cdots + \xi_{k+1})]$ a.s. since $n_0 \geq 0$ (a.s.). Choosing $f = \mathbf{1}_{(-\infty,a]}$, $a \geq 0$, we finally obtain

$$\forall k \geq 0, \quad \mathbb{P}_0\{n_{k+1} \leq a, \tau > k, T_k < S\} \leq \mathbb{Q}\{\xi_1 + \cdots + \xi_{k+1} \leq a\} \quad \text{(a.s.)}.$$

Step 2b. Deviation inequality. Choosing $a = (k+1)/2$ and applying Lemma 3.7 below, there exists $\delta_1 > 0$, only depending on C (and thus on d and Λ), such that, for $\delta \in (0, \delta_1)$,

$$\forall k \geq 0, \quad \mathbb{P}_0\{n_{k+1} \leq (k+1)/2, \tau > k, T_k < S\} \leq \delta^{(k+1)/2}. \tag{3.13}$$

Noting that $\{\tau > k+1\} \subset \{\tau > k\}$ and $\{T_{k+1} < S\} \subset \{T_k < S\}$, we also have for $\delta \in (0, \delta_1)$

$$\forall k \geq 1, \quad \mathbb{P}_0\{n_k \leq k/2, \tau > k, T_k < S\} \leq \delta^{k/2}. \tag{3.14}$$

In what follows, we assume $\delta \in (0, \delta_1)$ so that (3.13) and (3.14) hold.

Step 3. Exit time. We now evaluate the exit time S. For $0 \leq k < \tau$ and $T_k \leq t < T_{k+1}$, (3.9) yields

$$\begin{aligned}\|X_t - x_\infty\| = \Big\| &\frac{T_{k+1} - t}{T_{k+1} - T_k}(X_{T_k} - x_\infty) \\ &+ \delta^{-2n_k} \int_{T_k}^t (\psi(\lambda(X_s) - 1)(x_\infty - X_{T_k}) ds + \int_{T_k}^t \sigma(X_s) dW_s \Big\|.\end{aligned}$$

Therefore, following (3.12), for $0 \leq k < \tau$ and $T_k < S$,

$$\begin{aligned}&\exists t \in (T_k, T_{k+1}] : \|X_t - x_\infty\| \geq 3/4 \\ &\Rightarrow \sup_{T_k \leq t \leq T_{k+1}} \Big[\delta^{-n_k} \int_{T_k}^t \mathbf{1}_{\{\lambda(X_s) > \alpha/2\}} ds + \Big\|\int_{T_k}^t \sigma(X_s) dW_s\Big\|\Big] \geq 3/4 - \|X_{T_k} - x_\infty\|,\end{aligned} \tag{3.15}$$

so that (have in mind $\|\cdot\| \le |\cdot|$)

$$S \le T_{k+1} \Rightarrow \sup_{T_k \le t \le T_{k+1}} \Big[\delta^{-n_k} \int_{T_k}^t \mathbf{1}_{\{\lambda(X_s)>\alpha/2\}} ds + \Big| \int_{T_k}^t \sigma(X_s) dW_s \Big|\Big] \ge (3/4 - \delta^{n_k})^+.$$

By (3.12) and by Doob's maximal inequality, for $k \ge 0$, we have on the event $\{T_{k\wedge\tau} < S\}$ (of course, $T_{k\wedge\tau}$ is a stopping time)

$$\mathbb{P}_{T_{k\wedge\tau}}\{S \le T_{(k+1)\wedge\tau}\} \le C\delta^{2n_k+4}[(3/4 - \delta^{n_k})^+]^{-2} \mathbf{1}_{\{k<\tau\}}. \tag{3.16}$$

(With the same C as above, i.e. $C = 4/\alpha + 2d\Lambda$.) Setting $\nu_k = \delta^{2n_k+4}[(3/4 - \delta^{n_k})^+]^{-2}\mathbf{1}_{\{k<\tau\}}$ $(+\infty \cdot 0 = 0)$, we deduce

$$\mathbb{P}_0\{S > T_{(k+1)\wedge\tau}\} \ge \mathbb{E}_0[(1 - C\nu_k)^+ \mathbf{1}_{\{S > T_{k\wedge\tau}\}}]. \tag{3.17}$$

Since $\delta < 1/4$, we deduce from (3.14) that, for any $k \ge 1$ (below, use that $\nu_k \ne 0 \Rightarrow k < \tau$, use also that $n_k > k/2$ implies $\nu_k \le 4\delta^{k+4}$)

$$\begin{aligned} \mathbb{P}_0\{\nu_k > 4\delta^{k+4}, T_{k\wedge\tau} < S\} &= \mathbb{P}_0\{\nu_k > 4\delta^{k+4}, \tau > k, T_k < S\} \\ &\le \mathbb{P}_0\{n_k \le k/2, \tau > k, T_k < S\} \le \delta^{k/2}. \end{aligned} \tag{3.18}$$

Plugging (3.18) into (3.17), we have $\mathbb{P}_0\{S > T_{(k+1)\wedge\tau}\} \ge (1 - 4C\delta^{k+4})^+ \mathbb{P}_0\{S > T_{k\wedge\tau}\} - \delta^{k/2}$. By induction, we deduce that, for any $k \ge 1$, $\mathbb{P}_0\{S > T_{k\wedge\tau}\} \ge \prod_{i=1}^{k-1}(1 - 4C\delta^{i+4})^+ \mathbb{P}_0\{S > T_{1\wedge\tau}\} - \sum_{i=1}^{k-1} \delta^{i/2} \ge \prod_{i=1}^{+\infty}(1 - 4C\delta^{i+4})^+ \mathbb{P}_0\{S > T_{1\wedge\tau}\} - o(1)$. (Remind that $o(1)$ is purely deterministic.) Following (3.15) and (3.16) (with the bound $\|X_0 - x_\infty\| \le 1/4$), $\mathbb{P}_0\{S > T_{1\wedge\tau}\} \ge 1 - C(3/4 - \|X_0 - x_\infty\|)^{-2}\delta^4 \ge 1 - 4C\delta^4$. We deduce that, for any $k \ge 0$, $\mathbb{P}_0\{S > T_{k\wedge\tau}\} \ge \prod_{i=0}^{+\infty}(1 - 4C\delta^{i+4})^+ - o(1) = 1 - o(1)$. Letting $k \to +\infty$, we obtain:

$$\mathbb{P}_0\Big(\bigcap_{k\ge 0}\{S > T_{k\wedge\tau}\}\Big) \ge 1 - o(1), \text{ i.e. } \mathbb{P}_0\Big\{\exists k \ge 0 : S \le T_{k\wedge\tau}\Big\} \le o(1). \tag{3.19}$$

We are then able to get rid of the event $\{T_k < S\}$ in (3.13). Summing (3.13) over $k \ge 0$, we obtain $\mathbb{P}_0\{\exists 0 \le k < \tau : n_{k+1} \le (k+1)/2, T_k < S\} \le o(1)$. In light of (3.19), we deduce $\mathbb{P}_0\{\exists 0 \le k < \tau : n_{k+1} \le (k+1)/2\} \le o(1)$, so that

$$\mathbb{P}_0\Big(\bigcap_{k\ge 0}\{n_{k\wedge\tau} \ge (k\wedge\tau)/2\}\Big) \ge 1 - o(1). \tag{3.20}$$

(In fact, we have added the case $k = 0$ in the above intersection. This just follows from the relationship $n_0 \ge 0$.)

Step 5. Conclusion We now complete the proof. To this end, we set $R = 1 + \sum_{k\ge 0} \delta^k = 1 + 1/(1-\delta)$. This will be the "$R_0$" appearing in the final statement.

The idea is the following. With high probability (see (3.19)), the exit time is greater than all the times $(T_{k\wedge\tau})_{k\ge 0}$ so that the exit phenomenon can be forgotten. Now, if τ is infinite, then the process $(X_t)_{t\geq 0}$ converges towards $x_\infty \in V$ in time less than R on the event $\{\forall k \geq 0, n_k \geq k/2\}$. If τ is finite, there are two cases. If $\tau = \tau_1$, then $X_{T_\tau} = x_\infty$ and the process hits x_∞ in time less than R on the event $\{\forall k \geq 0, n_k \geq k/2\}$. If $\tau = \tau_2$, then the process hits V with a non-zero conditional probability between T_{τ_2} and T_{τ_2+1} under the action of the noise. Again, the hitting time is less than R on the event $\{\forall k \geq 0, n_k \geq k/2\}$.

Step 5a. Case a: $\tau = +\infty$. On the event $\{\tau = +\infty\} \cap \{\forall k \geq 0, n_k \geq k/2, T_k < S\}$, we have, for any $k \geq 0$, $T_k \leq \sum_{\ell=0}^{k-1} \delta^{2n_\ell} < R - 1$. Therefore, the non-decreasing sequence $(T_k)_{k\geq 0}$ converges towards some finite real T_∞. It is clear that $T_\infty \leq S$ and $T_\infty \leq R - 1$. Moreover, for any $k \geq 0$, $|X_{T_k} - x_\infty| \leq \delta^{n_k} \leq \delta^{k/2}$, so that $X_{T_\infty} = x_\infty \in V$. In particular, $T_\infty < S$. We deduce (with $T_\tau = T_\infty$ on $\{\tau = +\infty\}$)

$$\begin{aligned} &\mathbb{P}_0\{T_V \leq T_\tau < R \wedge S, \tau = +\infty\} \\ &\geq \mathbb{P}_0\Big(\{\tau = +\infty\} \cap \{\forall k \geq 0, n_{k\wedge\tau} \geq (k \wedge \tau)/2, T_{k\wedge\tau} < S\}\Big). \end{aligned} \tag{3.21}$$

Step 5b. Case b: $\tau = \tau_1 < +\infty$. The argument is almost the same as above. On the event $\{\tau < +\infty\} \cap \{\forall k \in \{0, \ldots, \tau\}, n_k \geq k/2, T_k < S\}$, $T_\tau \leq \sum_{k=0}^{\tau-1} \delta^{2n_k} < R$. Moreover, on the event $\{\tau = \tau_1 < +\infty\}$, $X_{T_\tau} = X_{T_{\tau_1}} = x_\infty \in V$. Therefore,

$$\begin{aligned} &\mathbb{P}_0\{T_V \leq T_\tau < R \wedge S, \tau = \tau_1 < +\infty\} \\ &\geq \mathbb{P}_0\Big(\{\tau = \tau_1 < +\infty\} \cap \{\forall k \geq 0, n_{k\wedge\tau} \geq (k \wedge \tau)/2, T_{k\wedge\tau} < S\}\Big). \end{aligned} \tag{3.22}$$

Step 5c. Case c: $\tau < +\infty$, $\tau_2 < \tau_1$. We start as above. On the event $\{\tau < +\infty\} \cap \{\forall k \in \{0, \ldots, \tau\}, n_k \geq k/2, T_k < S\}$, $T_\tau \leq T_{\tau+1} = \sum_{k=0}^{\tau} \delta^{2n_k} < R$. Therefore,

$$\begin{aligned} &\mathbb{P}_0\{T_V < T_{\tau+1} \wedge S \leq R \wedge S, \tau < +\infty, \tau_2 < \tau_1\} \\ &\geq \mathbb{P}_0\Big(\{\tau < +\infty, \tau_2 < \tau_1\} \cap \{\forall k \geq 0, n_{k\wedge\tau} \geq (k \wedge \tau)/2, T_{k\wedge\tau} < S\} \\ &\qquad \cap \{\exists t \in (T_\tau, T_{\tau+1}) : X_t \in V, \forall s \in (T_\tau, t], X_s \in Q(x_\infty, 3/4)\}\Big) \end{aligned} \tag{3.23}$$

We now apply Proposition 3.1. On the event $\{\tau < +\infty, \tau_2 < \tau_1\} \cap \{\forall k \geq 0, n_{k\wedge\tau} \geq (k \wedge \tau)/2, T_{k\wedge\tau} < S\}$ (which is in $\mathcal{F}_{T_\tau}$), we deduce from (3.8):

$$\mathbb{E}_{T_\tau}\Big[\int_{T_\tau}^{T_{\tau+1}} \lambda(X_s)ds\Big] \geq \delta^{2n_\tau+4}. \tag{3.24}$$

We also have $\|X_{T_\tau} - x_\infty\| < \delta^{n_\tau}$ (see (3.10)). In the specific case when $\tau = 0$ and $n_\tau = 0$, the bound $\|X_{T_\tau} - x_\infty\| = \|X_0 - x_\infty\| < 1/4$ is more useful. Moreover, by (3.9), the drift $(b_t)_{T_\tau \leq t \leq T_{\tau+1}}$ is bounded by $(T_{\tau+1} - T_\tau)^{-1}\|X_{T_\tau} - x_\infty\| \leq \delta^{-n_\tau}$. If $\tau = 0$ and $n_\tau = 0$, $(\|b_t\|)_{T_\tau \leq t \leq T_{\tau+1}}$ is bounded by 1 and thus by 4.

Therefore, we can apply Proposition 3.1 with $\rho = \rho_\tau$, where $\rho_\tau = \delta^{n_\tau}$ for $n_\tau \geq 1$ (on the above event, n_τ is always greater than 1 when $\tau \geq 1$) and $\rho_\tau = 1/4$

for $n_0 = 0$ and $\tau = 0$. In light of (3.24), we choose $\eta = \delta^4$ (when $\tau = 0$ and $n_0 = 0$, $\delta^{2n_\tau} = 1 \geq 1/16 = \rho^2$). We note from (3.7) that $|Q(x_\infty, 3\rho_\tau) \setminus V| \leq 3^d K_0 |Q_1 \setminus V|^{1/2} \rho_\tau^d$. Therefore, with μ and ε given by Proposition 3.1, we have, for $3^d K_0 |Q_1 \setminus V|^{1/2} \leq \mu(\delta^4)$,

$$\mathbb{P}_{T_\tau}\Big\{\exists t \in (T_\tau, T_{\tau+1}) : X_t \in V, \forall s \in (T_\tau, t], X_s \in Q(x_\infty, 3\rho_\tau)\Big\} \geq \varepsilon(\delta^4)$$

on the event $\{\tau < +\infty, \tau_2 < \tau_1\} \cap \{\forall k \geq 0, n_{k\wedge\tau} \geq (k \wedge \tau)/2, T_{k\wedge\tau} < S\}$. Since $\delta < 1/4$, ρ_τ is always less than $1/4$ on the event $\{\tau < +\infty, \tau_2 < \tau_1\} \cap \{\forall k \geq 0, n_{k\wedge\tau} \geq (k \wedge \tau)/2, T_{k\wedge\tau} < S\}$, so that $Q(x_\infty, 3\rho_\tau) \subset Q(x_\infty, 3/4)$. By (3.23), we finally obtain, for $3^d K_0 |Q_1 \setminus V|^{1/2} \leq \mu(\delta^4)$,

$$\begin{aligned}
&\mathbb{P}_0\{T_V < T_{\tau+1} \wedge S \leq R \wedge S, \tau < +\infty, \tau_2 < \tau_1\} \\
&\geq \varepsilon(\delta^4)\mathbb{P}_0\Big(\{\tau < +\infty, \tau_2 < \tau_1\} \cap \{\forall k \geq 0, n_{k\wedge\tau} \geq (k \wedge \tau)/2, T_{k\wedge\tau} < S\}\Big).
\end{aligned} \tag{3.25}$$

Step 5d. Putting cases a, b and c together. By (3.21), (3.22) and (3.25), we have $\mathbb{P}_0\{T_V < R \wedge S, T_V \leq T_{\tau+1}\} \geq \varepsilon(\delta^4)\mathbb{P}_0\{\forall k \geq 0, n_{k\wedge\tau} \geq (k \wedge \tau)/2, T_{k\wedge\tau} < S\}$ for $3^d K_0 |Q_1 \setminus V|^{1/2} \leq \mu(\delta^4)$. By (3.19) and (3.20), we can choose δ small enough such that the probability $\mathbb{P}_0\{\forall k \geq 0, n_{k\wedge\tau} \geq (k \wedge \tau)/2, T_{k\wedge\tau} < S\}$ is greater than $1/2$.

Step 6. Integrability of the drift. We choose δ and V as above. By the previous step, we know that the process X hits V before $T_{\tau+1} \wedge S$ with a non-zero probability. Therefore, we can kill the drift after $T_{\tau+1} \wedge S$. It is thus enough to prove the integrability of $(b_t \mathbf{1}_{\{t < T_{\tau+1}\wedge S\}})_{t\geq 0}$. By (3.9) and (3.10), we have, for any $p \in [1, 2)$,

$$\begin{aligned}
\mathbb{E}_0\Big[\int_0^{T_{\tau+1}\wedge S} \|b_t\|^p dt\Big] &\leq \sum_{k\geq 0} \mathbb{E}_0\Big[\mathbf{1}_{\{k\leq\tau, T_k < S\}} \int_{T_k}^{T_{k+1}} \|b_t\|^p dt\Big] \\
&\leq \sum_{k\geq 0} \mathbb{E}_0\Big[\delta^{(2-p)n_k} \mathbf{1}_{\{k\leq\tau, T_k<S\}}\Big] \quad (b_t = 0 \text{ for } t \geq T_{\tau_1}) \\
&\leq 1 + \sum_{k\geq 0} \mathbb{E}_0\Big[\delta^{(2-p)n_k} \mathbf{1}_{\{k<\tau, T_k<S\}}\Big] \quad (T_k < S \Rightarrow n_k \geq 0).
\end{aligned}$$

Now, for any $k \geq 0$, $\mathbb{E}_0[\delta^{(2-p)n_k}\mathbf{1}_{\{k<\tau,T_k<S\}}] \leq \delta^{(2-p)k/2} + \mathbb{P}_0\{n_k \leq k/2, k < \tau, T_k < S\}$. By (3.14), $\mathbb{P}_{T_0}\{n_k \leq k/2, k < \tau, T_k \leq S\} \leq \delta^{k/2}$. Setting $\Gamma_p = 1 + 1/(1 - \delta^{(2-p)/2}) + 1/(1 - \delta^{1/2})$, this proves that $\mathbb{E}_0 \int_0^{T_{\tau+1}\wedge S} \|b_t\|^p dt \leq \Gamma_p$.

Step 7. Proof of the deviation inequality. It remains to prove

Lemma 3.7 *Let C be a positive real and $(\xi_k)_{k\geq 1}$ a sequence of I.I.D. random variables with values in $\mathbb{Z}$ such that, for any $k \geq 1$,*

$$\mathbb{Q}\{\xi_k = -\ell\} = C\delta^{2(1+\ell)},\ \ell \geq 0, \quad \mathbb{Q}\{\xi_k = 1\} = 1 - C\delta^2/(1 - \delta^2),$$

on some probability space $(\Xi, \mathcal{A}, \mathbb{Q})$. (With δ small enough so that the law is well defined.) Then, there exists $\delta_0 > 0$, only depending on C, such that for $\delta \in (0, \delta_0)$, $\mathbb{Q}\Big\{\xi_1 + \cdots + \xi_k \leq k/2\Big\} \leq \delta^{k/2}$.

Proof. It is well seen that $\mathbb{E}^{\mathbb{Q}}[\xi_1] = 1 - o(1) > 1/2$ for δ small enough. It is thus enough to bound from below the Cramer transform H of ξ_1 given by $H(t) = \sup_{\lambda \in \mathbb{R}}[\lambda t - \ln(\phi(\lambda))]$, $t \in \mathbb{R}$, with $\phi(\lambda) = \mathbb{E}^{\mathbb{Q}}[\exp(\lambda \xi_1)]$. For $\lambda > \ln(\delta^2)$, $\phi(\lambda) = C\delta^2/(1 - \exp(-\lambda)\delta^2) + \exp(\lambda)[1 - C\delta^2/(1-\delta^2)]$. Choosing $\lambda = \ln(2\delta^2)$, $\phi(\ln(2\delta^2)) = [2(C+1) + o(1)]\delta^2$. Hence, for $t = 1/2$, $H(t) \geq -(1/2)\ln(\delta^2) - \ln(2(C+1) + o(1))$. For δ less than some $\delta_0 > 0$, we obtain $H(t) \geq -(1/4)\ln(\delta^2)$ and $t < \mathbb{E}^{\mathbb{Q}}[\xi_1]$, so that $\mathbb{Q}\{\xi_1 + \cdots + \xi_k \leq kt\} \leq \exp(-kH(t)) = \exp(k\ln(\delta^2)/4) = \delta^{k/2}$. □

4 Attainability of Small Sets

4.1 Attainability of a Small Ball

Following the standard Krylov and Safonov proof, we first prove that we can force the process X by an additional drift to let it hit a ball of small radius.

Lemma 4.1 *Keep the assumptions and notations of Proposition 3.6. Then, for any $\beta \in (0,1)$, there exist positive constants $\zeta(\beta)$, $r(\beta)$ and $(\gamma_p(\beta))_{p \in [1,2)}$ only depending on d, α, μ and Λ, such that, for any hypercube Q_1 of $\mathbb{R}^d$ of radius 1, any square integrable $\mathcal{F}_0$-measurable random variable X_0 with values in $\mathbb{R}^d$, any $\rho \in (0,1)$ and any $z \in Q_{1-\rho\beta}$ (hypercube of same center as Q_1 but of radius $1 - \rho\beta$), we can find a d-dimensional progressively-measurable process $(b_t)_{t \geq 0}$ such that $(b_t)_{t\geq 0}$ together with the process $(X_t)_{t \geq 0}$ equal to $\mathcal{S}_{X_0}((b_t)_{t \geq 0}, \sigma)$*

fulfill
$$\begin{cases} \forall t \geq 0, \quad \lambda(X_t) \geq \alpha \Rightarrow b_t = 0, \\ \forall p \in [1,2), \quad \mathbb{E}_0\Big[\displaystyle\int_0^{+\infty} |b_t|^p dt\Big] \leq \gamma_p(\beta)\rho^{2-p}, \end{cases}$$

and $\mathbb{P}_0\{T_{Q(z,\rho\beta)} < (r(\beta)\rho^2) \wedge S_{Q_1}\} \geq \zeta(\beta)$ a.e. on the event $\{X_0 \in Q_{1-\rho\beta}, \|X_0 - z\| \leq \rho\}$. ($T_{Q(z,\rho\beta)}$ is the first hitting time of the hypercube $Q(z,\rho\beta)$ and S_{Q_1} the first exit time from the hypercube Q_1 by X.)

Proof. As already explained in the proof of Proposition 3.6, we can assume that $X_0 \in Q_{1-\rho\beta} \cap \overline{Q}(z,\rho)$ (a.s.). It is also sufficient to perform the proof for small values of β. As in the proof of Proposition 3.6, we consider a smooth function ψ with values in $[0,1]$, matching 1 on $(-\infty, \alpha/2]$ and 0 on $[\alpha, +\infty)$. We set $b_t = \rho^{-2}\beta^{-3}(z - X_0)\mathbf{1}_{[0,\rho^2\beta^3]}(t)$, $t \geq 0$, and we consider $(X_t)_{t \geq 0}$ solution of

$$X_t = X_0 + \int_0^t \psi(\lambda(X_s))b_s ds + \int_0^t \sigma(X_s)dW_s, \quad t \geq 0.$$

The L^p bounds, $1 \leq p < 2$, of the drift easily follow. It thus remains to bound from below the probability of hitting $Q(z, \rho\beta)$. Define to this end $\mathbb{Q}$ as the probability on $(\Omega, \mathcal{F})$ admitting

$$Z = \exp\Big(\int_0^{\rho^2\beta^3} (1-\psi(\lambda(X_s)))\langle \sigma^{-1}(X_s)b_s, dW_s - \frac{1-\psi(\lambda(X_s))}{2}\sigma^{-1}(X_s)b_s ds\rangle\Big)$$

as density with respect to $\mathbb{P}$. We emphasize that the inverse of σ is well defined when λ is non-zero: we have $|\sigma^{-1}(X_s)b_s|^2 \leq \Lambda\rho^{-4}\beta^{-6}|X_0 - z|^2\lambda^{-1}(X_s) \leq 2d\Lambda\rho^{-2}\beta^{-6}\alpha^{-1}$ when $\lambda(X_s) \geq \alpha/2$. Setting $\mathbb{Q}_0 = \mathbb{Q}[\cdot|\mathcal{F}_0]$, we have, for any event $A \in \mathcal{F}$, $\mathbb{Q}_0(A) \leq \mathbb{E}_0(Z^2)^{1/2}\mathbb{P}_0^{1/2}(A) \leq \exp(2d\Lambda\beta^{-3}\alpha^{-1})\mathbb{P}_0^{1/2}(A)$ (a.s.). It is thus sufficient to prove the result under $\mathbb{Q}$. By Girsanov's theorem, the process $(\hat{W}_t = W_t - \int_0^{t\wedge(\rho^2\beta^3)}(1-\psi(\lambda(X_s)))\sigma^{-1}(X_s)b_s ds)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion under $\mathbb{Q}$. We write

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma(X_s)d\hat{W}_s, \quad t \geq 0.$$

Hence, $X_{\rho^2\beta^3} = z + \int_0^{\rho^2\beta^3}\sigma(X_s)d\hat{W}_s$, so that

$$\mathbb{Q}_0\{\|X_{\rho^2\beta^3} - z\| \geq \rho\beta\} \leq \rho^{-2}\beta^{-2}\mathbb{E}_0^{\mathbb{Q}}[|X_{\rho^2\beta^3} - z|^2] \leq d\Lambda\beta. \tag{4.1}$$

Now, for any $t \in [0, \rho^2\beta^3]$, $X_t = (1 - t\rho^{-2}\beta^{-3})X_0 + t\rho^{-2}\beta^{-3}z + \int_0^t \sigma(X_s)d\hat{W}_s$, so that $\|X_t\| \leq 1 - \rho\beta + \|\int_0^t \sigma(X_s)d\hat{W}_s\|$. By Doob's maximal inequality

$$\mathbb{Q}_0\{S_{Q_1} \leq \rho^2\beta^3\} \leq \mathbb{Q}_0\Big\{\sup_{0\leq t\leq \rho^2\beta^3}\Big\|\int_0^t \sigma(X_s)d\hat{W}_s\Big\| \geq \rho\beta\Big\} \leq d\Lambda\beta. \tag{4.2}$$

By (4.1) and (4.2), we deduce that $\mathbb{Q}_0\{S_{Q_1} > \rho^2\beta^3, \|X_{\rho^2\beta^3} - z\| < \rho\beta\} \geq 1 - 2d\Lambda\beta$ (a.s.). For $\beta < 1/(4d\Lambda)$, this is greater than $1/2$, so that $\mathbb{Q}_0\{T_{Q(z,\rho\beta)} < (\beta^3\rho^2) \wedge S_{Q_1}\} \geq 1/2$. □

4.2 Proof of the Main Result

We are now in position to complete the proof of Theorem 1.2. Again, we can assume ρ to be equal to 1. We then follow the original proof by Krylov and Safonov. To this end, we remind the reader of the following lemma of measure theory (see [3, Prop (7.2)]). (We adopt the same convention as in [3]: if Q is an open hypercube with z as center and $\rho > 0$ as radius, then $\hat{Q}$ denotes the closed hypercube with z center and 3ρ as radius).

Lemma 4.2 *Let $(q, \nu) \in (0,1)^2$. If $V \subset Q_1$ and $|V| \leq q|Q_1|$, then there exists a finite family $(C_i)_{i\in\mathcal{I}}$ of pairwise disjoint open hypercubes, all included in Q_1,* *such that: (1) for each $i \in \mathcal{I}$, $|V \cap C_i| > q|C_i|$, (2) $|V| \leq q(|D \cap Q_1| + \nu)$ with $D = \bigcup_{i\in\mathcal{I}} \hat{C}_i$.*

(Pay attention, in [3, Prop (7.2)], $Q(0,1)$ stands for an hypercube of volume 1.) The statement of [3, Prop (7.2)] is in fact slightly different. In [3, Prop (7.2)], the family $\mathcal{I}$ is countable so that ν can be chosen equal to zero, that is $|V| \leq q|D \cap Q_1|$. We thus show that Lemma 4.2 with $\mathcal{I}$ countable and $\nu = 0$ implies our own version of Lemma 4.2: by choosing $(\mathcal{I}_n)_{n\geq 1}$ a sequence of increasing sets of indices such that $\#(\mathcal{I}_n) = n$, $n \geq 1$, and $\cup_{n\geq 1}\mathcal{I}_n = \mathcal{I}$, we clearly have $|D \cap Q_1| \leq |D_n \cap Q_1| + \sum_{i\notin \mathcal{I}_n} |\hat{C}_i|$, $D_n = \cup_{i\in\mathcal{I}_n}\hat{C}_i$, $n \geq 1$. By choosing n large enough, we obtain the desired result. (Of course, the sum $\sum_{i\in\mathcal{I}} |\hat{C}_i|$ is finite since $\sum_{i\in\mathcal{I}} |\hat{C}_i| \leq 3^d \sum_{i\in\mathcal{I}} |C_i| \leq 3^d|Q_1| < +\infty$.) The advantage is the following: in our own version of Lemma 4.2, D is closed.

Step 1. Initialization. Proposition 3.6 says that Theorem 1.2 holds true for $\mu \leq \mu_0$. We now establish Theorem 1.2 for $|Q_1 \setminus V|/|Q_1| \in (\mu_0, \mu_1]$, with $\mu_1 = \mu_0(1+(1-\mu_0)^2)$. (It is clear that the mapping $x \in (0,1) \mapsto x(1+(1-x)^2)$ is an increasing mapping from $(0,1)$ onto itself and that it is above $x \in (0,1) \mapsto x$.) We then apply Lemma 4.2 with $q = 1-\mu_0$ and $\nu = (\mu_0^2|Q_1|/2) \vee (1/2) \in (0,1)$. For the resulting D, we have $|D \cap Q_1| \geq |V|/q - \nu \geq [(1-\mu_1)/(1-\mu_0) - \mu_0^2/2]|Q_1| = (1 - \mu_0 + \mu_0^2/2)|Q_1|$.

Set now $E = D \cap \overline{Q}_{(1-\mu_0^2/2)^{1/d}} \subset Q_1$. Then $|E| \geq |D \cap Q_1| + |Q_{(1-\mu_0^2/2)^{1/d}}| - |Q_1| \geq (1-\mu_0)|Q_1|$. By Proposition 3.6, we can find a d-dimensional $(\mathcal{F}_t)_{t\geq 0}$ progressively measurable process $(b_t^0)_{t\geq 0}$ such that

$$\begin{Bmatrix} \forall t \geq 0, \quad \lambda(X_t^0) \geq \alpha \Rightarrow b_t^0 = 0 \\ \forall p \in [1,2), \quad \mathbb{E}_0 \displaystyle\int_0^{+\infty} |b_t^0|^p dt \leq \Gamma_p \end{Bmatrix} \quad \text{and} \quad \mathbb{P}_0\{T_E^0 < R_0 \wedge S_{Q_1}^0\} \geq \varepsilon_0,$$

where T_E^0 is the first hitting time of E and $S_{Q_1}^0$ the first exit time from Q_1 by $(X_t^0)_{t\geq 0}$, equal to $\mathcal{S}_{X_0}((b_t^0)_{t\geq 0}, \sigma)$. (Here, X_0 is some $\mathcal{F}_0$-measurable random variable with values in $Q_{1/8}$: it may be x_0 as in the statement of Theorem 1.2.)

Step 2. Hitting intermediate sets. We define $\tau_0 = T_E^0 \wedge R_0 \wedge S_{Q_1}^0$. If $T_E^0 < R_0 \wedge S_{Q_1}^0$, then $X_{\tau_0}^0$ belongs to E since E is closed. In particular, there exists $i \in \mathcal{I}$ such that $X_{\tau_0}^0$ belongs to $\hat{C}_i$. We then denote by x_i the center of C_i and by s_i its radius: $\hat{C}_i$ is the closed hypercube with x_i as center and $3s_i$ as radius, so that $\|X_{\tau_0}^0 - x_i\| \leq 3s_i$. Since $X_{\tau_0}^0$ belongs to E, we also have $X_{\tau_0}^0 \in \overline{Q}_{(1-\mu_0^2/2)^{1/d}} \subset Q_{1-(1-(1-\mu_0^2/2)^{1/d})s_i}$ since $s_i \in (0,1)$. Setting $\rho_i = 3s_i$ and $\beta = \min([1-(1-\mu_0^2/2)^{1/d}]/3, 1/48)$, we have $\|X_{\tau_0}^0 - x_i\| \leq \rho_i$, $X_{\tau_0}^0 \in Q_{1-\rho_i\beta}$ and $x_i \in Q_{1-\rho_i\beta}$ since $Q(x_i, \rho_i\beta) \subset Q(x_i, s_i) = C_i \subset Q_1$. (The term $1/48$ may be explained as follows: $\rho_i\beta \leq s_i/16$ so that $\overline{Q}(x_i, \rho_i\beta) \subset Q(x_i, s_i/8)$. The factor $1/8$ is the same as in Proposition 3.6.) Following the proof of [3, Thm (7.4)], we denote by C_i^* the closed hypercube of center x_i and radius $s_i/16$.

We now apply Lemma 4.1 to $z = x_i$, $\rho = \rho_i$ and β as above, $i \in \mathcal{I}$. For any $i \in \mathcal{I}$, we can find a d-dimensional $(\mathcal{F}_{\tau_0+t})_{t\geq 0}$ progressively-measurable process $(b_t^{1,i})_{t\geq 0}$ such that, $(b_t^{1,i})_{t\geq 0}$ together with $(X_t^{1,i})_{t\geq 0}$, solution of the SDE

$$X_t^{1,i} = X_{\tau_0}^0 + \int_0^t b_s^{1,i} ds + \int_0^t \sigma(X_s^{1,i}) dW_s^1, \quad W_t^1 = W_{\tau_0+t} - W_{\tau_0}, \quad t \geq 0,$$

fulfill

$$\begin{cases} \forall t \geq 0, \quad \lambda(X_t^{1,i}) \geq \alpha \Rightarrow b_t^{1,i} = 0, \\ \forall p \in [1,2), \quad \mathbb{E}_{\tau_0}\Big[\int_0^{+\infty} |b_t^{1,i}|^p dt\Big] \leq \gamma_p(\beta)\rho_i^{2-p}, \end{cases} \tag{4.3}$$

and

$$\mathbb{P}_{\tau_0}\{T_{C_i^*}^{1,i} < (\rho_i^2 r(\beta)) \wedge S_{Q_1}^{1,i}\} \geq \zeta(\beta) \tag{4.4}$$

a.e. on the event $\{X_{\tau_0}^0 \in E \cap \hat{C}_i\}$, where $T_{C_i^*}^{1,i}$, $i \in \mathcal{I}$, is the first hitting time of the hypercube C_i^* and $S_{Q_1}^{1,i}$ the first exit time from the hypercube Q_1 by $(X_{t,i}^1)_{t\geq 0}$. Without loss of generality, we can assume that $\mathcal{I}$ is a subset of $\mathbb{N}^*$. Therefore, at time τ_0, we can define $I_0 = \inf\{i \in \mathcal{I} : X_{\tau_0}^0 \in \hat{C}_i\}$ (with $I_0 = +\infty$ if $X_{\tau_0}^0 \notin \bigcup_{i\in\mathcal{I}} \hat{C}_i$). We then set:

$$\forall t \geq 0, \quad b_t^1 = \sum_{i\in\mathcal{I}} \mathbf{1}_{\{I_0=i\}} b_t^{1,i}, \quad X_t^1 = \sum_{i\in\mathcal{I}\cup\{+\infty\}} \mathbf{1}_{\{I_0=i\}} X_t^{1,i},$$

where $(X_t^{1,+\infty})_{t\geq 0}$ is the solution of the SDE $dX_t^{1,+\infty} = \sigma(X_t^{1,+\infty})dW_t^1$, $t \geq 0$, $X_0^{1,+\infty} = X_{\tau_0}^0$. It is clear that $(X_t^1)_{t\geq 0}$ is $(\mathcal{F}_{\tau_0+t})_{t\geq 0}$ progressively measurable and solves the SDE

$$X_t^1 = X_{\tau_0}^0 + \int_0^t b_s^1 ds + \int_0^t \sigma(X_s^1) dW_s^1, \quad t \geq 0,$$

and that the pair $(b_t^1, X_t^1)_{t\geq 0}$ fulfills (4.3) with $\gamma_p(\beta)3^{2-p}$ instead of $\gamma_p(\beta)\rho_i^{2-p}$. (In short, $\rho_i \leq 3$ for every $i \in \mathcal{I}$.) Moreover, setting $F = \bigcup_{i\in\mathcal{I}} C_i^*$, we deduce from (4.4) that $\mathbb{P}_{\tau_0}\{T_F^1 < (9r(\beta)) \wedge S_{Q_1}^1\} \geq \zeta(\beta)$ a.e. on $\{X_{\tau_0}^0 \in E\}$, T_F^1 being the first hitting time of F and $S_{Q_1}^1$ the first exit time from Q_1 by $(X_t^1)_{t\geq 0}$.

Step 3. Hitting the prescribed Borel subset. We set $\tau_1 = T_F^1 \wedge (9r(\beta)) \wedge S_{Q_1}^1$. It is a stopping time for the filtration $(\mathcal{F}_{\tau_0+t})_{t\geq 0}$. If $\tau_1^1 < (9r(\beta)) \wedge S_{Q_1}^1$, then $X_{\tau_1}^1$ belongs to F and thus to some $Q(x_i, s_i/8)$, $i \in \mathcal{I}$. For each $i \in \mathcal{I}$, $|V \cap Q(x_i, s_i)| = |V \cap C_i| > (1-\mu_0)|Q(x_i, s_i)|$, so that we can apply Proposition 3.6. For any $i \in \mathcal{I}$, we can find a d-dimensional $(\mathcal{F}_{\tau_0+\tau_1+t})_{t\geq 0}$ progressively-measurable process $(b_t^{2,i})_{t\geq 0}$ such that $(b_t^{2,i})_{t\geq 0}$ together with $(X_t^{2,i})_{t\geq 0}$, solution of the SDE

$$X_t^{2,i} = X_{\tau_1}^1 + \int_0^t b_s^{2,i} ds + \int_0^t \sigma(X_s^{2,i}) dW_s^2, \quad W_t^2 = W_{\tau_0+\tau_1+t} - W_{\tau_0+\tau_1}, \quad t \geq 0,$$

fulfill

$$\begin{cases} \forall t \geq 0, \quad \lambda(X_t^{2,i}) \geq \alpha \Rightarrow b_t^{2,i} = 0, \\ \forall p \in [1,2), \quad \mathbb{E}\int_0^{+\infty} |b_t^{2,i}|^p dt \leq \Gamma_p s_i^{2-p}, \end{cases} \tag{4.5}$$

and

$$\mathbb{P}_{\tau_0+\tau_1}\{T_V^{2,i} < (R_0 s_i^2) \wedge S_{C_i}^{2,i}\} \geq \varepsilon_0 \tag{4.6}$$

a.e. on the event $\{X_{\tau_1}^1 \in Q(x_i, s_i/8)\}$, where $T_V^{2,i}$ is the first hitting time of V and $S_{C_i}^{2,i}$ the first exit time from the hypercube C_i by the process $(X_t^{2,i})_{t\geq 0}$. At time τ_1, we can define $I_1 = \inf\{i \in \mathcal{I} : X_{\tau_1}^1 \in C_i^*\}$ (with $I_1 = +\infty$ if $X_{\tau_1}^1 \notin \bigcup_{i\in\mathcal{I}} C_i^*$). Following Step 2, we set

$$\forall t \geq 0, \quad b_t^2 = \sum_{i\in\mathcal{I}} \mathbf{1}_{\{I_1=i\}} b_t^{2,i}, \quad X_t^2 = \sum_{i\in\mathcal{I}\cup\{+\infty\}} \mathbf{1}_{\{I_1=i\}} X_t^{2,i},$$

where $(X_t^{2,+\infty})_{t\geq 0}$ is the solution of the SDE $dX_t^{2,+\infty} = \sigma(X_t^{2,+\infty})dW_t^2$, $t \geq 0$, $X_0^{2,+\infty} = X_{\tau_1}^1$. As above, $(X_t^2)_{t\geq 0}$ is $(\mathcal{F}_{\tau_0+\tau_1+t})_{t\geq 0}$ progressively measurable and solves the SDE

$$X_t^2 = X_{\tau_1}^1 + \int_0^t b_s^2 ds + \int_0^t \sigma(X_s^2)dW_s^2, \quad t \geq 0.$$

Moreover, the pair $(b_t^2, X_t^2)_{t\geq 0}$ fulfills (4.5) with $\Gamma_p s_i^{2-p}$ replaced by Γ_p. (Indeed, $s_i \leq 1$.) By (4.6), we also have $\mathbb{P}_{\tau_0+\tau_1}(\{T_V^2 < R_0 \wedge S_{Q_1}^2\}) \geq \varepsilon_0$ a.e. on $\{X_{\tau_1}^1 \in F\}$, where T_V^2 is the first hitting time of V and $S_{Q_1}^2$ is the first exit time from Q_1 by $(X_t^2)_{t\geq 0}$.

Step 4. Conclusion. We finally define:

$$b_t = b_t^0 \mathbf{1}_{\{0\leq t<\tau^0\}} + b_{t+\tau^0}^1 \mathbf{1}_{\{\tau^0\leq t<\tau^0+\tau^1\}} + b_{t+\tau^0+\tau^1}^2 \mathbf{1}_{\{t\geq\tau^0+\tau^1\}}, \quad t \geq 0.$$

We define $(X_t)_{t\geq 0}$ as $\mathcal{S}_{X_0}((b_t)_{t\geq 0}, \sigma)$. Then, $\mathbb{P}_0\{T_V < (2R_0 + 9r(\beta)) \wedge S_{Q_1}\} \geq \varepsilon_0^2 \zeta(\beta)$, where T_V is the first hitting time of V and S_{Q_1} the first exit time from Q_1 by $(X_t)_{t\geq 0}$. The integrability of b is easily checked as well as the vanishing property (i.e. $b_t = 0$ if $\lambda(X_t) \geq \alpha$, $t \geq 0$). This proves that Theorem 1.2 holds true for $|Q_1 \setminus V| \leq \mu_1$. By induction, we can prove that it holds true for $|Q_1 \setminus V| \leq \mu_n$, $n \geq 0$, where $(\mu_n)_{n\geq 0}$ is the sequence given by $\mu_{n+1} = \mu_n(1+(1-\mu_n)^2)$. (We emphasize that, for each $n \geq 1$, we can apply Proposition 3.6 with μ_n instead of μ_0 since X_0 is chosen random in the above demonstration.) The sequence $(\mu_n)_{n\geq 0}$ is non-decreasing. Since $\mu_0 > 0$, the limit $\mu_\infty = \lim_{n\to+\infty} \mu_n$ is clearly equal to 1. In other words, we can reach any real $\mu \in (0,1)$, such that $|Q_1 \setminus V| \leq \mu$, in a finite number of iterations. □

References

[1] Barles, G., Busca, J. Existence and comparison results for fully nonlinear degenerate elliptic equations without zeroth-order term. Comm. Partial Differential Equations 26 (2001), 2323–2337.

[2] Barles, G., Da Lio, F. Local $C^{0,\alpha}$ estimates for viscosity solutions of Neumann-type boundary value problems. J. Differential Equations 225 (2006), 202–241.

[3] Bass, R. F. Diffusions and elliptic operators. Springer-Verlag, New York, 1998.

[4] Caffarelli, L. A. Interior estimates for fully nonlinear elliptic equations, Ann. of Math. 130 (1989) 189213.

[5] Caffarelli, L. A., Cabré, X. Fully Nonlinear Elliptic Equations, in: Amer. Math. Soc. Collog. Publ., vol. 43, American Mathematical Society, Providence, RI, 1995.

[6] Crandall, M. G., Kocan, M., Soravia, P., Swiech, A. On the equivalence of various weak notions of solutions of elliptic PDEs with measurable ingredients. Progress in elliptic and parabolic partial differential equations (Capri, 1994), 136–162, Pitman Res. Notes Math. Ser., 350, Longman, Harlow, 1996.

[7] Crandall, M. G., Ishii, H., Lions, P.-L. Uniqueness of viscosity solutions of Hamilton Jacobi equations revisited. J. Math. Soc. Japan 39 (1987), 581–596.

[8] Delarue, F. Estimates of the solutions of a system of quasilinear PDEs. A probabilistic scheme. Séminaire de Probabilités, XXXVII, pp. 290-332.

[9] De Giorgi, E. Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. Mem. Accad. Sci. Torino Cl. Sci. Fis. Mat. Natur. 3 (1957), 25–43.

[10] DiBenedetto, E., Urbano, J. M., Vespri, V. Current issues on singular and degenerate evolution equations. Evolutionary equations. Vol. I, 169–286, Handb. Differ. Equ., North-Holland, Amsterdam, 2004.

[11] Gilbarg, D., Trudinger, N. S. Elliptic partial differential equations of second order. Second edition. Springer-Verlag, Berlin, 1983.

[12] Ishii, H., Lions, P.-L. Viscosity solutions of fully nonlinear second-order elliptic partial differential equations. J. Differential Equations 83 (1990), 26–78.

[13] Jakobsen, E. R., Karlsen, K. H. Continuous dependence estimates for viscosity solutions of fully nonlinear degenerate elliptic equations. Electron. J. Differential Equations 39 (2002), 10 pp.

[14] Jensen, R. Uniqueness criteria for viscosity solutions of fully nonlinear elliptic partial differential equations. Indiana Univ. Math. J. 38 (1989), no. 3, 629–667.

[15] Katsoulakis, M. A. A representation formula and regularizing properties for viscosity solutions of second-order fully nonlinear degenerate parabolic equations. Nonlinear Anal. 24 (1995), 147–158.

[16] Krylov, N. V. Controlled diffusion processes. Translated from the Russian by A. B. Aries. Springer-Verlag, New York-Berlin, 1980.

[17] Krylov, N. V., Safonov, M. V. An estimate for the probability of a diffusion process hitting a set of positive measure. Dokl. Akad. Nauk SSSR 245 (1979), 18–20.

[18] Krylov, N. V., Safonov, M. V. A property of the solutions of parabolic equations with measurable coefficients. (Russian) Izv. Akad. Nauk SSSR Ser. Mat. 44 (1980), 161–175.

[19] Ladyzhenskaya, O. A., Solonnikov, V. A., Ural'ceva, N. N. Linear and quasilinear equations of parabolic type. American Mathematical Society, Providence, R.I. 1967.

[20] Ladyzhenskaya, O. A., Ural'ceva, N. N. Estimate of the Hlder norm of solutions of second-order elliptic equations of general form. Boundary value problems of mathematical physics and related questions in the theory of functions, 12. Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) 96 (1980), 161–168, 309.

[21] Lasry, J.-M., Lions, P.-L. A remark on regularization in Hilbert spaces. Israel J. Math. 55 (1986), 257–266.

[22] Nash, J. Continuity of solutions of parabolic and elliptic equations. Amer. J. Math. 80 (1958), 931–954.

[23] Pardoux, É., Peng, S. G. Backward stochastic differential equations and quasilinear parabolic partial differential equations. Stochastic partial differential equations and their applications (Charlotte, NC, 1991), 200–217, Lecture Notes in Control and Inform. Sci., 176, Springer, Berlin, 1992.

[24] Pucci, P. Serrin, J. The maximum principle. Progress in Nonlinear Differential Equations and their Applications, 73. Birkhäuser Verlag, Basel, 2007.

[25] Serrin, J. Local behavior of solutions of quasi-linear equations. Acta Math. 111 (1964), 247–302.

[26] Stroock, D. W., Varadhan, S. R. S. Multidimensional diffusion processes. Springer-Verlag, Berlin-New York, 1979.